AD-A098 055 NATIONAL RESEARCH COUNCIL WASHINGTON DC COMMITTEE ON --ETC F/G 5/2
NATIONAL NEEDS FOR CRITICALLY EVALUATED PHYSICAL AND CHEMICAL D--ETC(U)
1978

UNCLASSIFIED NL

1 OF 1
AD A
098055

END
DATE
FILMED
5-81
DTIC

LEVEL II

AD A098055

National Needs for Critically Evaluated Physical and Chemical Data

DTIC
ELECTE
APR 22 1981
D

Committee on Data Needs

Numerical Data Advisory Board

Assembly of Mathematical and Physical Sciences

DTIC FILE COPY

DISTRIBUTION STATEMENT A
Approved for public release;
Distribution Unlimited

National Needs for Critically Evaluated Physical and Chemical Data.

(11) 1978 / (12) 99

Committee on Data Needs
Numerical Data Advisory Board
Assembly of Mathematical and Physical Sciences
National Research Council

NATIONAL ACADEMY OF SCIENCES
Washington, D.C. 1978

NOTICE The project that is the subject of this report was approved by the Governing Board of the National Research Council, whose members are drawn from the Councils of the National Academy of Sciences, the National Academy of Engineering, and the Institute of Medicine. The members of the Committee responsible for the report were chosen for their special competences and with regard for appropriate balance.

This report has been reviewed by a group other than the authors according to procedures approved by a Report Review Committee consisting of members of the National Academy of Sciences, the National Academy of Engineering, and the Institute of Medicine.

Available from
Numerical Data Advisory Board
2101 Constitution Avenue
Washington, D.C. 20418

Committee on Data Needs

Walter H. Stockmayer, *Chairman*
Dartmouth College

Edward P. Bartkus
E.I. du Pont de Nemours & Company

Carlos M. Bowman
The Dow Chemical Company

Leo Brewer
University of California, Berkeley

Joseph J. Katz
Argonne National Laboratory

John L. Margrave
Rice University

John W. Murdock
Informatics, Inc.

Richard F. Taschek
Los Alamos Scientific Laboratory

Kurt L. Wray
Physical Sciences, Inc.

Numerical Data Advisory Board

PREFACE

This report concerns a problem that is crucial to the scientific and technological efforts of the nation, namely, to determine the level of critical evaluation needed to meet our national needs for reliable scientific data. The report deals specifically with the critical evaluation of numerical data, a special aspect of the more general question of communication of information in an evermore complex society. This is not a new problem, and indeed it has been studied a number of times before. Invariably such studies have concluded that the activity was grossly underfunded. To date, however, little budgetary action has resulted.

The suggestion of still another study in this area originated in a letter from Sidney Benson (then at Stanford Research Institute and now at the University of Southern California) to National Academy of Sciences President Philip Handler in 1973. Its planning was encouraged by the National Science Foundation (NSF), its eventual sponsor. After strong support by the Numerical Data Advisory Board of the National Research Council, the proposed study was endorsed by the Executive Committee of the Assembly of Mathematical and Physical Sciences (AMPS) in March 1975, funding for the study was secured from NSF, and the present committee was appointed to carry it out.

Three objectives of the Committee's work, as spelled out in the proposal to NSF were (1) to survey the organizations that currently carry out critical data evaluation and determine the present level of funding for such activities; (2) to study the role of organized data collections and the required depth of evaluation of the data in previous R&D programs and assess the benefits of such activities relative to their cost; (3) to identify current and future data needs in major national R&D programs, particularly those concerned with energy, environmental quality, and materials utilization.

Aside from bringing knowledge of existing activities and needs up to date, as in the first and last objectives, the study was planned to improve on its precursors, insofar as possible, by the inclusion of more specific information, notably in respect to the cost/benefit investigations of particular data evaluation programs.

Although this report addresses only U.S. activities, it should be borne in mind that a number of other countries--particularly Germany, the United Kingdom, Japan, and the Soviet Union--have substantial government-supported data programs. Furthermore, several organizations, such as the International Atomic Energy Agency and the ICSU Committee on Data for Science and Technology (CODATA), play an active role. Current U.S. data programs appear to be well coordinated with efforts in other countries. It is important that this coordination continue, so that U.S. scientists and engineers can take maximum advantage of work done elsewhere.

The organization of the Committee's work was greatly assisted by William Spindel, Executive Secretary of the Office of Chemistry and Chemical Technology, National Research Council, and Hendrik van Olphen, Executive Secretary of the Numerical Data Advisory Board until March 1977. Everett Johnson, Consultant, gathered most of the detailed information and prepared the initial draft report. Robert S. Marvin, Executive Secretary of the Numerical Data Advisory Board from April 1977 through February 1978, is responsible for the organization and most of the writing of the final report.

Walter H. Stockmayer, *Chairman*
Committee on Data Needs

CONTENTS

1
Summary and Major Recommendations

1. SUMMARY

Reliable values of numerical data that express in quantitative terms the properties and behavior of materials are essential in all branches of science and technology and are needed to arrive at valid decisions whenever a governmental or industrial decision involves elements of science and technology. The scientific literature contains many valuable data covering a wide range of diverse fields. Unfortunately, it also contains many erroneous values. A substantial intellectual effort is required to select reliable values from the large and growing total of those reported, (see Section 3.1).

The selection of the best available values for data in a given field requires the background of a specialist in that field. Most users are not specialists in all the fields in which they require data. Furthermore it is inefficient for many individuals who need the same data for different purposes to each go through this selection process.

For this reason, a number of specialized data centers have been established to compile and evaluate data in a systematic fashion. Typically, such a center gathers all the data applicable to its limited area, assesses the validity of the measurements on which these data are based, selects recommended or best values, and attempts to estimate how far the "true" values are from those recommended. These results are then published and made available to all who need them. (see Chapter ~~4)~~ 7

The cost of this evaluation in established data centers is a fraction of 1 percent of the cost of obtaining the original data, ~~(see~~ Chapter 5). The benefits to the nation of having compilations of reliable data readily available are substantial. Such compilations save time for engineers and scientists in research and development. If the reliability of a needed set of data is known, designs can be made more precise, tolerances reduced, and R&D options narrowed. The resulting savings can amount each year to from one to several thousand times the cost of evaluation (see Chapter 6).

The present level of data evaluation activities is about one third to one half that needed to carry out activities planned for the next five years by federal agencies with major mission responsibilities that require the use of reliable scientific data (see Chapter 7). These same data will be used by industry also, but the benefits of evaluated data are spread among so many users that the major responsibility for financing their acquisition must rest with the federal government (see Chapter 8).

1.2 MAJOR RECOMMENDATIONS

Our three major recommendations follow:

1. The present annual support for organized data evaluation activities of slightly under $7 million should be increased over a period of five years to $18 million. For reasons outlined in Chapter 8 this support will have to come primarily from the federal government.
2. When a particular mission relies heavily on results from a field of research, responsibility for data compilation and evaluation in that field should be accepted by the agency responsible for the mission. The Office of Standard Reference Data of the National Bureau of Standards should be responsible for categories of data of very broad utility and for general coordination of the overall system.
3. Each agency should be required to place its responsibility for data compilation and evaluation on one key official at a level high enough to ensure that the agency's responsibilities in this area will be fulfilled.

Additional recommendations appear in Chapter 8.

2
INTRODUCTION

The measurement of the inherent properties of substances and materials has been a key element in the progress of science and the translation of scientific understanding into useful technology. The data that result from these measurements represent a resource that can be used for a variety of applications, often over a long time span. In today's world, heavily dependent on science and technology, it is important to understand the extent to which time and money are lost as a result of the lack of reliable data needed by scientists and engineers for the accomplishment of their work.

Unreliable data can be worse than no data at all. Their use can lead to poorly conceived experiments, ineffective or inefficient manufacturing plants, and a waste of both effort and resources. To those studying basic science or practicing its applications in technology, engineering, and industry *assurance of reliability of the data base is indispensable.*

The importance of such assurance became apparent in the early days of the modern age, when a consensus developed among astronomers that the data of Tycho Brahe exceeded in accuracy and precision all of that accumulated in the previous millennium. This gave Kepler the confidence needed to devise a model of the solar system consistent with Brahe's data and provide the world with an entry into modern times.

Today, the quality of data may have more immediate effects. This was clearly demonstrated during the development of fission reactors, which in turn were crucial to nuclear-weapon development. Initially the possibility that graphite reactors could be made to work depended entirely on the capture cross section of neutrons in carbon. Until it was confirmed that this neutron absorption cross section was quantitatively small enough so that graphite could be used as a moderator for reactors, no support could be committed for a major development project. Fermi's 1940 measurement of the carbon capture cross section provided the assurance necessary for the support, leading soon thereafter to the Hanford plutonium production reactors.

Another option for use as a moderator was deuterium, but here the U.S. measurement program was inadequate to guarantee the success of a heavy-water reactor. In Germany, an erroneous measurement of the carbon cross section, and moreover an *erroneous evaluation* of this measurement as dependable, effectively stopped their reactor program because the other alternative--heavy water--required a much more extensive industrial effort.

During the course of the Manhattan Project, significant efforts were expended on the evaluation of various nuclear constants. Some of the more critical ones were the fission cross sections, the number of neutrons per fission, and the capture cross sections of certain fission product isotopes, such as ^{135}Xe. The importance of evaluation lay in the fact that often several measurements by different methods differed by amounts highly critical to design, or in some cases even feasibility, of nuclear-weapon or reactor systems.

In contrast to the above example dealing with an urgent national need and having a well-defined path from science to application, the importance of evaluated data to *scientific* understanding can be exemplified in Maria Goeppert-Mayer's elucidation of the shell model of the nucleus, which led to her sharing the Nobel Prize in 1963 with Jensen. In this case, the intimations that the nucleus had a well-defined shell structure had been studied inconclusively for nearly two decades by many scientists. The large masses of data of different kinds in part only confused matters, but as systematic evaluation proceeded, it was possible to place more and more confidence in the precision and reliability of the many thousands of data values and, finally, to produce a sound theory.

It should be noted that this was accomplished in a short period of time, at least in part because of the systematic data evaluation that either existed or was done by Goeppert-Mayer herself; statements (perhaps apocryphal) to this effect are often attributed directly to her. Just a hundred years earlier, a similar situation with respect to *atomic* structure instead of *nuclear* structure led Mendeleyev (and Lothar Meyer) to propose the periodic table of the chemical elements. Again this was possible only because of the systematic evaluation of a large body of often inadequate and discrepant data to determine its quantitative reliability and therefore applicability.

These examples are illustrative of the power of having a reliable data base in science and technology. Its credibility must be established by systematic professional evaluation of the initial raw information derived from experiment. Such major codifications and integrations of knowledge as those described for nuclear and atomic structure not only benefit fundamental science but establish a capability for quantitative interpolation, extrapolation, and new directions of understanding that are particularly effective when applied to technology and engineering.

The interdisciplinary nature of much of modern technology is self-evident. Solid-state physicists require data generated by crystallographers, metallurgists, chemists, and other physicists. Nuclear engineers must have not only nuclear-physics data but also nuclear-chemistry, mechanical-property, solid-state, and metallurgical data. Credible, reliable data immediately at hand optimize the utilization of scientific knowledge for technological purposes.

This report attempts to place in perspective the cost of generating physical and chemical data through laboratory measurements and of evaluating such data and prepa ing compilations of reliable data that are readily accessible to a wide variety of users. An estimate will be made of the benefits to the nation that accrue from support of such compilation and evaluation activities, and, finally a projection of future needs for reliable data in support of major national programs will be made.

3

Survey of Data Evaluation Activity

3.1 CRITICALLY EVALUATED DATA

In this report we discuss critically evaluated data, sometimes termed standard reference data. "Data" as used here means the quantitative results of scientific measurements that can be reproduced at other locations and times.

Our attention will be focused on data that represent inherent material properties. Such data rarely are numbers read directly from a laboratory instrument. Rather, such numbers must be used in a calculation, based on theory, to obtain the value for the material property sought and more often than not require the use of other numbers not measured in the experiment, such as a density, a molecular weight, or a value for the acceleration of gravity. A part of the process of critical evaluation involves checking the report of the work to make sure that both the appropriate theory and the best values of the various constants required have been used. Another part involves checking the description of the experimental arrangement used, to ensure that the temperature, pressure, and other ambient conditions were adequately controlled and that proper corrections were made where required.

Another aspect of critical evaluation is ascertaining that the sample of material measured was actually representative of the material of interest. Sometimes chemical purity alone is sufficient to ensure this, but often structural details of the sample affect the measured values. Thus, the requirements for a properly characterized sample depend on the property being measured. This characterization may require a specification of the sample history, for example, the thermal or mechanical treatment to which the sample has been subjected.

In addition, the reported property values can be checked against those of related properties of the same material. There are often relationships between various properties that must be satisfied. If they are not satisfied the evaluator must decide which measurement was in error. Based on such analyses of all the reported results, a recommended or "best" value is selected.

Finally, a full critical evaluation includes a quantitative assessment of the effect of various sources of error or uncertainty and gives a range about the recommended value within which the "true" value is expected to lie. One component of this uncertainty is the precision of the measurements, representing the reproducibility attainable. But the more important contribution is generally due to systematic errors, which include inherent limitations in the construction and calibration of instruments and (often unavoidable) deviations from the conditions assumed in the theory of the measurement. This is often a subtle and difficult quantity to evaluate, and at times it can only be estimated from a comparison of two quite different measurements of the same quantity.

The most familiar examples of this type of evaluation are those involving the fundamental constants such as the speed of light, for which the total estimated uncertainties are a few parts per million or less. But it is the quantitative statement of uncertainty, not its magnitude, that is significant here. An uncertainty of 10 percent may be quite acceptable for many purposes *if that uncertainty is known.* The problem is that for most of the data reported in the primary literature or tabulated in handbooks, there is no consistent attempt to estimate the uncertainty in the fashion described. It is quite common to find stated uncertainties based entirely on the precision of the measurements involved, and in such cases the difference between two measurements will often be ten times the stated uncertainty of either. A few examples will illustrate this point:

(a) L. J. Kieffer [*J. Chem. Documentation 9*, 167, (1969)] found that two independent measurements of the cross section for the ionization of atomic helium differed by 25 percent, ten times the uncertainty estimated by those making the measurements.

(b) Aksel A. Bothner-By (in *Advances in Magnetic Resonance*, Vol. I, J. S. Waugh, ed., Academic Press, New York, 1965), concluded that 90 percent of the high-resolution NMR coupling constant data published in the primary literature were so unreliable as to be not worth considering.

(c) H. J. M. Hanley and G. E. Childs [*Science 159*, 1114-1116 (8 March 1968)] concluded that the correct values for the viscosity of gases at 600 to 2000 K were 10 percent higher than those commonly accepted.

(d) R. W. Powell and Y. S. Touloukian [*Science 181*, 999-1008 (14 September 1973)], discussing the results of a critical evaluation of the thermal conductivity of the elements carried out at the Thermophysical Properties Research Center at Purdue University, pointed out that the values selected and those listed in a respected and widely used handbook differed by 18 percent or more for 22 elements. For 14 elements they differed by 30 percent or more.

3.2 SOURCES OF DATA

Most of the measurements of interest here are published in one of the established professional journals dealing with physics, chemistry, or engineering. We have no definite figures on the number of articles that contain data on material properties, but a high percentage will contain data, theory, or contributions to techniques of measurement that would be of concern to someone carrying out a critical evaluation of the type described above.

In 1976, *Chemical Abstracts* (C.A.) covered a total of 390,905 documents (not including patents) of which about 328,000 were journal articles, about 84,000 of them published in U.S. journals (figures from Russel Rowlett and Paul Swartzentruber of Chemical Abstracts Service). These figures probably include most, though certainly not all, articles that would be needed for a critical evaluation of data on some material property. On the other hand, there is a small but unknown fraction of these articles that would not be pertinent. Since we are interested only in an indication of the magnitude of the total problem, rather than a precise measure, we may assume that there are somewhat over 300,000 articles published in the world annually that have some bearing on the critical evaluation of material properties.

The above number will probably continue to grow, though at a slower rate than it did during the 1960's. In fact, the number of documents in C.A. decreased by 1329 from 1975 to 1976. King Research, Inc., has recently published *Statistical Indicators of Scientific and Technical Communication (1960-1980)*, 1977 edition, NTIS: PB 278-279 (Price Code A-16), in which they project the rates of growth of scholarly articles published in the United States in several fields of science. For the five-year period 1975-1980 they project a total growth in the number of articles in the physical sciences as 7 percent; in engineering, 19 percent; and in life sciences, 29 percent (see their Table 3.3). The rate of growth in some other countries has been greater in recent years. The growth in numbers of articles in C.A. from 1972 through 1976 was 17 percent, but the percentage from U.S. journals has declined steadily from 36.6 percent in 1951 to 25.8 percent in 1975. Since we must include articles published in other countries and some outside the physical science category of the King Survey, we assume that the number of articles of concern to us will continue to increase at an average rate of 2 percent per year.

The C.A. Collective Indexes show a total of 3,085,199 documents covered over the period 1967 through 1976. This would include 2,580,000 journal articles if the percentage has remained constant over this period. Since we find (see Chapter 7) that existing data centers have an average backlog of two years and cover less than half of the fields in which data evaluation is needed, it seems reasonable to assume that over 2 million of these articles remain to be searched and their data extracted and analyzed.

To summarize, we are concerned with about 300,000 published articles per year, increasing by about 2 percent a year, plus an existing backlog of over 2 million older articles.

3.3 EXAMPLES OF EVALUATED DATA

To illustrate both the scientific contributions and the practical importance of critical evaluation, we give examples covering two types of data, chemical reaction rate constants and thermal conductivity.

3.3.1 A Chemical Reaction Rate Constant

The reaction

$$CO + OH \rightarrow CO_2 + H$$

is a basic chain propagating mechanism in the combustion of all organic fuels and is the principal oxidizing reaction for carbon monoxide. It plays a significant role in the water-gas reaction, in air pollution, and in atmospheric chemistry. Reliable values of the rate constant and its temperature dependence are needed for atmospheric modeling, for incinerator design, and for many other industrial problems involving combustion. Figure 3.1 shows the values of this constant available in 1976 (taken, with the ommission of some details, from D. L. Baulch, D. D. Drysdale, J. Duxbury, and S. J. Grant, *Evaluated Kinetic Data for High Temperature Reactions*, Vol. 3, Buttersworths, London, 1976), with the solid line showing the values recommended on the basis of a critical evaluation.

The first problem here is to decide which of the grossly discordant values at temperatures below 500 K ($10^3\ T^{-1} \gtrsim 2$) should be used, and this requires detailed study of the various measurements by a specialist. Another problem facing a nonspecialist is the fact that standard reaction rate theory leads to the expectation of Arrhenius-type behavior, which would correspond to a single straight line on this plot. To fit these data in such a fashion one would probably use a represenation something like one of the two dashed lines shown (added to the figure from Baulch *et al.*). Indeed, lines quite similar to both of those shown had been proposed in earlier studies. Some of the differences shown are over two orders of magnitude (note the logarithmic scale on the ordinate), and extrapolation to lower temperatures would yield an even greater discrepancy. The low-temperature range is the one important in atmospheric modeling.

In 1972, J. E. Wilson, Jr. [*J. Phus. Chem. Ref. Data 1*(2), 535-574 (1972)] concluded that the extremely low values shown in Figure 3.1 were in error. Baulch *et al.* agreed and also utilized the results of a theoretical calculation published in 1971 which predicted that the rate of this reaction *should* show a non-Arrhenius behavior. These considerations, plus a detailed analysis of the other measurements shown, led to their recommended values shown by the solid line in Figure 3.1. This example illustrates (1) the need for critical evaluation of data by experts, both to select the "best" values and to eliminate those that are grossly in error; (2) the importance of considering the best theories for the behavior of a given property (most earlier evaluations, particularly those prior to 1972, had attempted to fit the available data with a single straight line, and were obviously unable to represent all the valid measurements); and (3) the need for periodic re-examination of previously evaluated properties, in the light of newer measurements and theories.

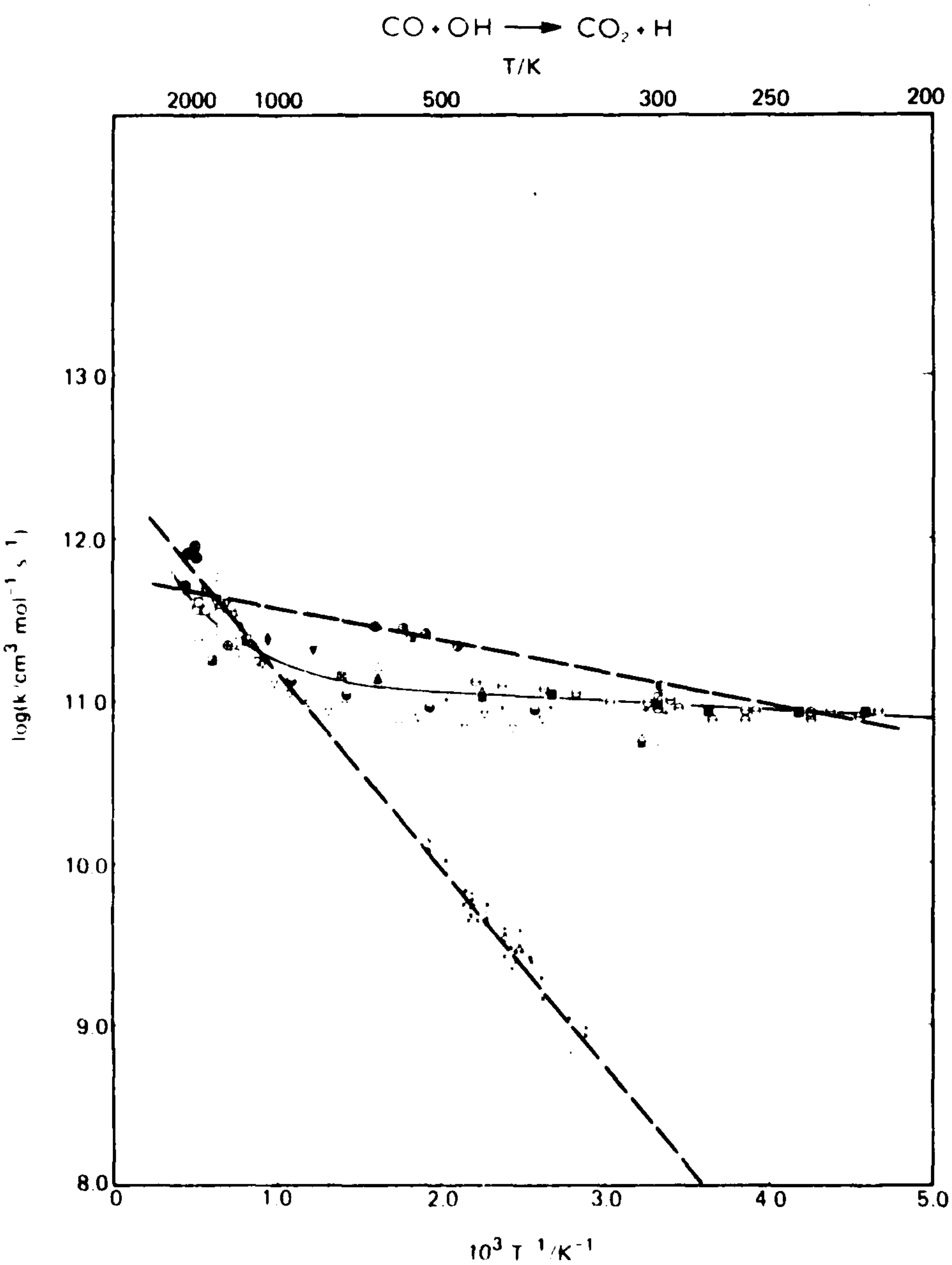

FIGURE 3.1 Rate constant for the reaction CO + OH → CO_2 + H. (From D. L. Baulch, D. D. Drysdale, J. Duxbury, and S. J. Grant, *Evaluated Kinetic Data for High Temperature Reactions*, Vol. 3, Butterworths, London, 1976. Some detail has been omitted for clarity, and the dashed lines have been added.)

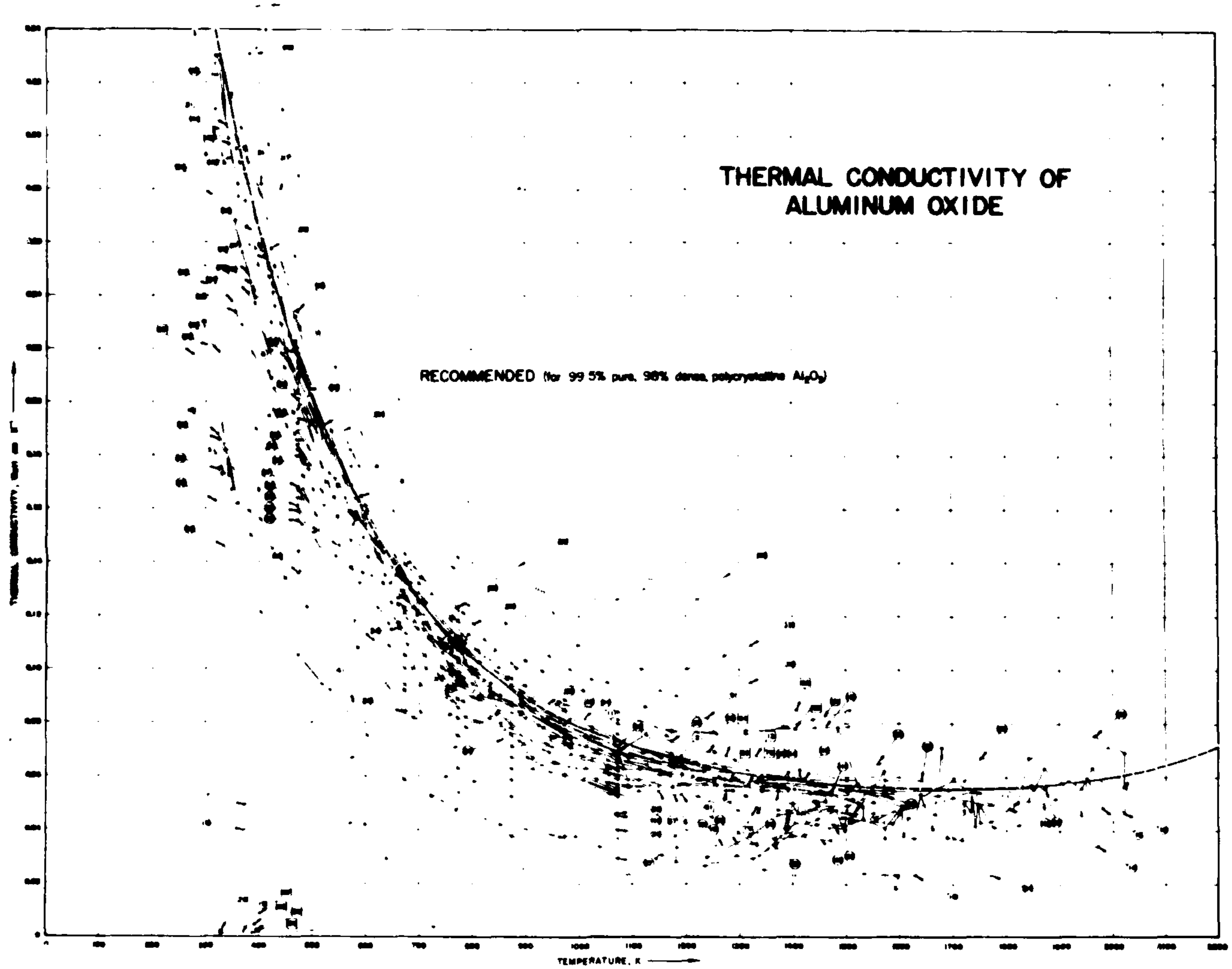

FIGURE 3.2 Thermal conductivity of aluminum oxide. (From R. W. Powell, C. Y. Ho, and P. E. Liley, NSRDS-NBS 8, 1966.)

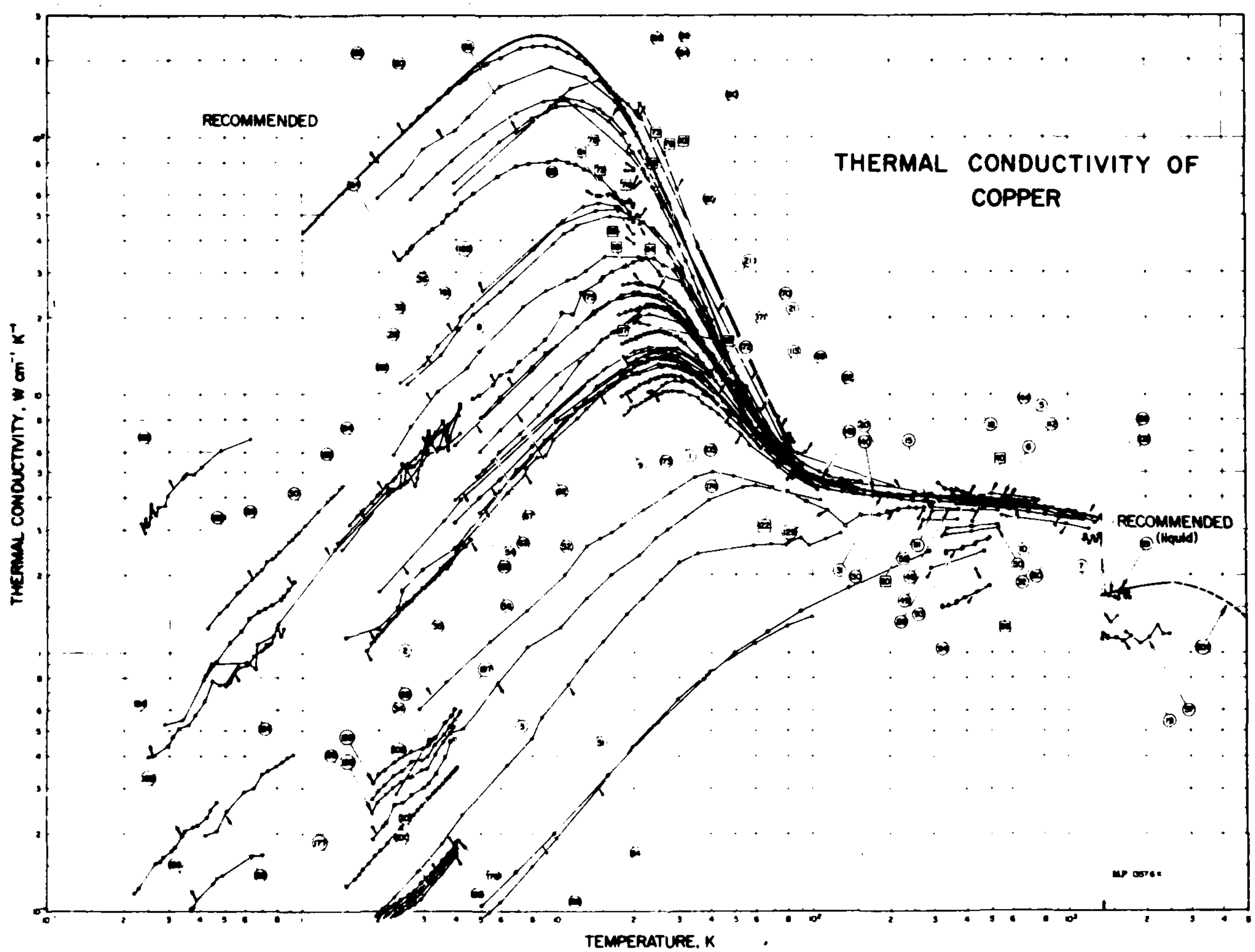

FIGURE 3.3 Thermal conductivity of copper. (From C. Y. Ho, R. W. Powell, and P. E. Liley, *J. Phys. Chem. Ref. Data 3*, Suppl. 1, 1974.)

3.3.2 Thermal Conductivity

Aluminum oxide is used to make radomes, missile nose cones, spark plug cores, valve seats, resistor cores, circuit breakers, electrical insulators, grinding wheels, crucibles, and many other items where its excellent abrasion, chemical and thermal shock resistance, thermal conductivity, mechanical strength, and dielectric characteristics are important. Obviously reliable values of thermal conductivity are required in the design of many of these parts.

Figure 3.2 shows a plot of the rather discrepant values found in the literature and the recommended values selected for a commercially available high-purity dense sample, based on the type of critical evaluation described earlier (R. W. Powell, C. Y. Ho, and P. E. Liley, *Thermal Conductivity of Selected Materials*, NSRDS-NBS 8, p. 73, 1966). Note that in this case the material characterization requires the specification of purity, density, and the fact that the material is polycrystalline. In much of the temperature range shown, the highest values shown are eight times the lowest, and a nonspecialist would certainly have a difficult time choosing a reliable value from those reported in the original literature.

To show that such divergence is not peculiar to aluminum oxide, we show in Figure 3.3 values for the thermal conductivity of copper, one of the oldest and most extensively studied metals. Here the spread in reported values is up to three orders of magnitude. (C. Y. Ho, R. W. Powell, and P. E. Liley, "Thermal Conductivity of the Elements," *J. Phys. Chem. Ref. Data 3*, Suppl. 1, 244, 1974).

These are not isolated examples. Similar cases can be found in every issue of the *Journal of Physical and Chemical Reference Data* and in the output of each of the data centers listed in Appendix A. They illustrate graphically the problems facing an engineer or scientist if he or she must rely on unevaluated research reports to find needed data on material properties and the enormous consolidation and rectification of diverse findings that is achieved by careful and systematic evaluation. We attempt a quantitative assessment of the benefits of such activities in Chapter 6.

4

DATA AND INFORMATION CENTERS

4.1 DATA CENTERS

The process of critical evaluation described above requires a substantial intellectual effort on the part of scientists with experience in the field concerned. Generally, the most competent evaluators are active participants in research on the materials and measurements involved. Until some 40 or 50 years ago, some of the leading figures in each field would devote a year or so at some stage in their careers to carrying out such evaluations in their specialty, much as they might on other occasions prepare a definitive review article. This was the general pattern followed in the compilations contained in the *International Critical Tables*. The growing volume of the literature, however, makes the maintenance of a comprehensive list of publications and extraction of the data a task that can be carried out effectively only on a continuing, long-term basis. For this reason, the last 30 years have seen the gradual development of a number of continuing data centers.

A data center is a more or less permanent organization that accepts the responsibility for accumulating the basic publications and other sources of numerical data on material properties within a specified area. It files and indexes these sources to permit ready retrieval of the numerical data. It also normally carries out the function of critical evaluation outlined above and may also permit its files to be used by others concerned with data evaluation in its field.

We have identified 37 such continuing data centers, listed in Appendix A. They vary in size from 1 to 26 professionals, with budgets ranging from $3000 to $1,200,000 per year. Two are supported by industrial groups, the others primarily by various branches of the federal government. Not all these centers perform the entire task of literature searching, data extraction, and critical evaluation. Some, like the Physical Data Group at Lawrence Livermore Laboratory, do not search the literature; their data are supplied by other centers both here and abroad, and they are primarily concerned with critical evaluation and preparation of convenient tables of data. Other groups, such as the National Nuclear Data Center, search and extract data from the literature and provide critically evaluated data; they also calculate and prepare tables of nuclear data in a form most suitable for those engaged in nuclear reactor design and operation. All 37 centers listed do carry out critical evaluation of data in their field.

Periodically, when the data covering a reasonable range of properties and/or materials are evaluated, a complete account is published, giving the original data and references, discussing shortcomings and strengths in the various measurements, and, finally, giving a set of recommended or "best" values with an estimate of the uncertainties in such values. Such a process represents an enormous consolidation and condensation of the original literature, as will be discussed in more quantitative terms in the next chapter.

The various data centers use other means of presenting their results to meet special needs. All of them maintain files of their evaluated data, and many have established computerized storage and retrieval schemes. From these files they frequently assemble tables of selected values for particular purposes, answer specific inquiries about data, and prepare various special publications addressed to those who need selected portions of the data. In some cases data are available on computer tapes, which can be purchased or leased.

Most of these data centers, as well as a number of short-term data projects, are part of the National Standard Reference Data System (NSRDS), which was established by the Federal Council on Science and Technology (FCST) in 1963. The FCST designated the National Bureau of Standards (NBS) as the focal point in the federal government for promoting and coordinating the critical evaluation of numerical data in the physical sciences. The "Standard Reference Data Act," Public Law 90-396, passed by Congress in 1968 further emphasized the central role of NBS in this System.

The Office of Standard Reference Data (OSRD), established at NBS in 1963, has the responsibility for allocating that part of the NBS budget that is spent on critical data evaluation, both within the NBS technical divisions and through contracts with outside groups. The staff members act as monitors for all projects supported by the Office. They maintain close contact with other data-compilation activities, both in the United States and abroad, and attempt to avoid needless duplication and to improve coverage of all important technical areas. The Numerical Data Advisory Board of the National Research Council was established to provide guidance to OSRD and to other federal agencies concerning their data problems; it carries out various studies to determine the status of the field, identify problem areas, and suggest solutions.

The major data publication, supervised and edited by the staff of OSRD, is the quarterly *Journal of Physical and Chemical Reference Data* (*JPCRD*), published jointly since 1972 by OSRD, the American Institute of Physics, and the American Chemical Society. By the end of 1977 (Vol. 6), it had published 109 articles in 6877 pages plus three supplements with a total of 1979 pages. In addition, OSRD has published another 78 papers in other NBS publication series or private journals and has prepared several data tapes. Individual articles from the *JPCRD* are also available for sale. All the *JPCRD* volumes and supplements are available on microfilm as well as in hard copy.

4.2 INFORMATION ANALYSIS CENTERS

The data activities discussed in this report constitute a subclass of a larger set of activities dealing with scientific and technical information. There are many organized centers that collect, analyze, and consolidate technical information of various types. The *Directory of Federally Supported Information Analysis Centers* (3rd edition, 1974), compiled by the National Referral Center of the Library of Congress (available from NTIS, ISBN 0-8444-0128-5), lists 108 such centers. These include most of the 37 centers listed in Appendix A but also include many others that do not deal with material properties or do not carry out data evaluation. For example, one compiles meteorological data, another compiles and analyzes data on mineral resources, and another on hearing, language, speech, and communication disorders. One center collects, evaluates, and disseminates information related to machining operations on all types of materials, another serves a similar role in the field of nondestructive testing. Though some of these centers collect and disseminate material properties data and others evaluate various types of information we have included in this study only centers that evaluate data on material properties.

5

Cost of Data Acquisition and Critical Evaluation

One of the chief objectives of this study is to establish both the costs and the benefits of critically evaluated data. The cost figures given in this chapter are for an operating data center that has already established its basic files and procedures for searching the literature in its field, identifying and procuring articles of interest, extracting the data, filing and indexing articles (and, in some cases, data from the articles) for ready retrieval, and evaluating the data. It would take much longer for an individual without established files to locate and extract data, and critical evaluation would probably take longer for someone without prior experience in evaluation, even though such an individual might have all the experimental and theoretical background required.

The 37 centers listed in Appendix A locate and add to their files a total of 65,000 documents per year. This represents only those documents selected as pertinent to their mission, as a result of examining a much larger number. The budgets of these 37 centers total $6,798,000 per year. Since most of the time and expense for an established center goes into evaluation, rather than locating and filing documents, the cost of locating, cataloging, and filing material in a fashion that permits ready retrieval for evaluation is only a fraction of the average of $100 per document required for the full process.

Of much greater significance is the average cost of evaluation of a group of data points that are presented as a unit. This "unit" will vary with the type of data considered; it may be a reaction rate or some other property as a function of temperature, pressure, or concentration. In other cases, it may be a tabulation of related properties, like various thermodynamic properties of a particular compound. In practice, the original presentations of evaluated data seem to be in terms of "units" that are reasonable to use for the comparisons made here.

Obviously, there is considerable variation in the number of primary references used to derive a unit of data and, consequently, in the cost of evaluation. Y. S. Touloukian, Director of the Center for Information and Numerical Data Analysis and Synthesis (CINDAS) estimates the cost of deriving a typical curve showing the thermal conductivity of aluminum as a function of temperature as $22,000 (private communication). This involved analysis of 71 references and covered a temperature range of over 900 kelvins.

David Garvin, Director of the Kinetic Data Center at NBS, estimates the average time required to evaluate a single reaction rate constant as a function of temperature as one man-week, varying from two days to four weeks (private communication). This involves the time of a PhD with at least five years' experience, which corresponds to a cost (including overhead) of about $1000.

The Joint Army Navy Air Force (JANAF) Thermochemical Tables data center effort has been in operation for approximately 16 years at a total cost of $2,300,000 [estimate by J. Masi at the Air Force Office of Scientific Research (AFOSR)] and has produced 2239 data sheets. This again represents a cost of about $1000 per data sheet (such a sheet serving as a unit in this case).

The Thermodynamic Data Center at NBS has been engaged since 1952 in the revision of NBS Circular 500, which covers thermodynamic data at one standard temperature on all elements, inorganic, and simple organic compounds. An estimate based on its funding and output since 1964 gives a figure of $800 per data sheet.

We shall use an estimate of $1000 as the average cost of evaluation and compilation of one unit of data. To put this cost in perspective, we now attempt to estimate the cost of the original research on which a typical evaluation was based. A count of the number of references and units of data in 12 papers from the *Journal of Physical and Chemical Reference Data* yields an average of 9.6 papers per unit, with a high of 27.3 and a low of 2.0. We shall use a rounded figure of 10 references for the average.

J. D. Frame and F. N. Narin (*Federation Proceedings 35*(14), 2529-2532, 1976), based on a study of the relationship of NIH grants to universities and the publications resulting from such grants, arrived at an estimated cost per paper of $31,000 in 1967 dollars. Based on the GNP Implicit Price Deflator (*Statistical Indicators, loc. cit.*, Table 2.12), this translates into $52,000 1976 dollars.

H. S. Milton, in *Cost of Research Index: 1920-65* (quoted by D. J. deSolla Price in *Science and Technology*, October 1967) gives a price tag of $20,000 for basic research papers in all fields worldwide. This would correspond to $34,000 in 1976 dollars.

An average of one paper per man-year is often used. For example, figures given by L. G. Burchinal (*Journal of Library Science, 14*(2), June 1977) and by A. H. Rosenfeld and P. R. Stevens (*Proc. 5th International CODATA Conference*, Pergamon Press, New York, 1977, pp. 19-23) yield about this number, and it is consistent with the experience in several national laboratories. Using a median annual salary in the physical sciences in 1976 of $22,600 (*Stat. Ind., loc. cit.*, Table 2.10), and allowing for 100 percent overhead, this gives a cost of $45,000 per paper.

We shall use the median of these figures, $45,000, as an estimate of the cost of research reported in a single paper used in evaluation of a unit of data. There is an additional publication cost per paper of about $400 (H. W. Koch, *Physics Today*, April 1968, pp. 41-49).

Thus, the data reported in ten papers representing an initial research cost of $450,000 and an additional publication cost of $4000 can be evaluated, summarized in one "unit" of data, and made readily accessible wherever needed for an additional expenditure of $1000. The cost of a comprehensive program of data collection and evaluation would be less than 0.2 percent of expenditures for basic research, since not all such research deals with material properties. And it is a much smaller fraction of total R&D expenses, since basic research accounts for only 13 percent of those costs (NSF 77-310, *National Patterns of R&D Resources; Funds and Manpower in the United States, 1953-1977*).

In the next chapter we demonstrate that the importance of this activity far outweighs its modest cost.

6

Benefits of Systematically Evaluated Data

Reliable data are required in all aspects of research and development and in the design of most products, industrial plants, and processes. They are also needed to assess the need for and impact of governmental programs and regulations concerned with all types of environmental, safety, and health questions.

6.1 GENERAL

In most cases data are needed on a wide range of properties and materials. At the request of the Committee on Data Needs, the Task Group for Scientific and Technical Information of the Industrial Research Institute, Inc. (IRI) conducted a survey of the data needs of 243 companies belonging to IRI. The 75 responses represented the categories of industrial chemical, electrical/electronic, packaging/paper, equipment/steel, food/drug, and personal care/home product companies. The 14 fields covered by the questionnaire were atomic structure, microstructure (two levels), thermodynamic, thermal, mechanical and acoustic, optical, electrical, magnetic, dielectric, nuclear (radiation damage), chemical and electrochemical, biological, and surface properties. Despite the diversity in categories and fields, some companies in each category indicated a need for data in each of the fields except for dielectric properties (one category) and radiation damage (two categories) (see Appendix B).

An earlier survey by the Materials Information Committee of the Federation of Materials Societies (quoted in *Materials Policy Handbook* prepared by the Science Policy Research Division, Congressional Research Service, Library of Congress, June 1977, Superintendent of Documents, 90-443) concluded that there is a broad recognition of the critical importance of materials information, including, but not limited to, data, and that improvements are needed in the evaluation, condensation, and presentation of such data.

The most obvious benefit of systematically evaluated and readily available compilations of recognized reliability is that their existence eliminates the need for repetitious searching of the literature by many workers to find and select the same values. The cost in each case is equal to or greater than that of having the job done once in an established center and then made available to all. In the case of the JANAF

Tables, mentioned in Chapter 5, which are incorporated into many computer files, the savings each year are several thousand times the cost of evaluation.

Conyers Herring (Appendix to *Report of the Task Group on the Economics of Primary Publication*, Committee on Scientific and Technical Communication, NAS-NAE, 1970, p. 119) concluded after a survey of several studies that research chemists spend an average of five hours or more a week reading the primary journals, with less time spent by research scientists in other fields and by engineers. If an average of one hour per week, or fifty hours per year, could be saved by having comprehensive compilations of reliable data available, this would amount to (50/2000) × \$45,000 = \$1125 per man per year. (The figure of \$45,000 is from *Stat. Ind., loc. cit.*, Table 2.1 + 100% overhead.) For the 82,600 physical scientists and engineers engaged in basic and applied research in 1974, this would come to \$93 million per year.

This type of calculation gives a narrow picture of the benefits of a data evaluation program, because much larger savings can come from the use of reliable data as compared with questionable data. If the reliability of the data is known, designs can be made more precise, tolerances reduced, and R&D options narrowed. The wasteful practice of overdesigning industrial plants to allow for uncertainties in the data can be minimized. This is particularly important today, since rising costs have led to strong pressures to eliminate the pilot plant and prototype development stages by which design parameters have traditionally been optimized. More and more decisions in all sectors of U.S. industry are being made on the basis of mathematical modeling and simulation, utilizing the capabilities of the modern digital computers.

The availability of reliable data bases for input to these models is translated into direct savings both in capital investment for plants and equipment and in operating costs. In addition, other important constraints such as minimizing energy consumption and avoiding the discharge of environmental pollutants require accurate data for input into the design programs.

Howard B. Hipkin, a senior engineer with the Bechtel Corporation, says (letter dated March 2, 1977) ". . . inadequate or unreliable data are reflected in excessive but undefined safety factors in design. While it is fairly easy to establish the money lost on a plant that does not work, it is virtually impossible to estimate the money lost on a plant that works well but is overdesigned." Even if we cannot quantify these savings, the total costs are so large that even a small percentage saved is significant.

Petroleum refineries are about 95 percent energy efficient, significantly better than most other industrial plants. An important contribution to reaching this level has come from American Petroleum Institute Research Project 44, started in 1942 and other data acquisition and evaluation projects sponsored both by API and various individual companies. Project 44 is now absorbed the Thermodynamics Research Center at Texas A&M University and has long been recognized as the source of much of the data needed by the industry.

There is another significant benefit that comes from the existence of a systematic program of data evaluation. An established data center can use its files and experience to prepare special compilations in a short time. This time saving can be extremely important when new regulations or legislation dealing with environmental, health, or safety problems are under consideration. In such cases it is also extremely important that the data used be recognized as reliable and from an authoritative source. The formal publication by data centers of their evaluated data, with full documentation, gives assurance that such values are the best available.

We have not attempted to set any total dollar figures on these benefits, but we give below some specific illustrations of savings in both money and time resulting from the availability of data in existing centers.

6.2 DESIGN OF NUCLEAR REACTORS

The National Nuclear Data Center at Brookhaven is an outgrowth of an effort by the Atomic Energy Commission to systematize data handling in the early days of atomic power. It compiles and evaluates data and also calculates various quantities used in the design of nuclear power plants, thus ensuring that various designs are based on a common set of data. One of the factors calculated, based on data for the radioactive decay of various fission products, has until recently had an associated uncertainty of 20 percent. Examination of the data showed the need for new measurements, which have now been made, permitting a reduction in the above uncertainty to 6 percent. This, in turn, results in a savings of at least $10 million per reactor (letter of March 22, 1977, from the Director, NNDC). This amounts to a saving of $710 million for the 71 reactors now under construction and a total saving of $1.5 billion if the additional 76 reactors with limited authorization or on order are included. (Numbers of reactors from *Status of U.S. Nuclear Electric Generating Capacity,* Divison of Nuclear Research and Applications, ERDA, March 1, 1977.)

Other important sets of figures from the National Nuclear Data Center are the cross sections for neutron capture by fission products. These affect reactor criticality and hence the amount of uranium-235 needed per reactor. A recent re-evaluation has reduced the uncertainty in these cross sections by 10 percent, for a saving of $2 million per reactor (letter from Director, NNDC, *loc. cit.*). This is a fuel saving of 5 percent, valued at $294 million for the 147 reactors under construction or planned.

These savings should be contrasted with the estimate [by P. B. Hemmig, Chief, Physics Branch, Division of Reactor Development Demonstration, ERDA (now part of the Department of Energy)] of $50 million spent on *all* nuclear data compilation activity since its inception about 1950.

6.3 DEVELOPMENT OF ROCKET FUELS

The JANAF Tables, mentioned in Chapter 5, were started in 1960 to meet the need for a central source of reliable thermochemical data for the development of high-performance rocket engines. The need was dramatized by the failure of an intensive effort in the early 1950's to develop rocket fuels containing boron compounds, which had looked promising on the basis of the original information available. One of the problems was that the then accepted value for the heat of formation of gaseous metaboric acid, HBO_2, was significantly in error, but this error was discovered only after the expenditure of considerable time and money. This was but one of many thermochemical values of uncertain reliability on which optimistic calculations and proposals for new fuels were based. A central source was needed to collect and evaluate existing data, select recommended values, assign uncertainties to these values, and point out cases where new measurements were required to resolve existing discrepancies.

Since 1960, the JANAF Tables have been prepared and published at an estimated cost through 1976 of $2.3 million, an average of about $140,000 per year. They are now incorporated into computer programs at all the major rocket research and development centers. They are also widely used in research on chemicals, explosives, lasers, and many other areas unrelated to the original motivation for the project.

In this example we want to concentrate on the potential savings that can be made if it is possible to screen out unworkable systems, like the boron fuels mentioned, on the basis of calculations using reliable data, without incurring the large costs required in testing a new fuel system. The procedure for testing a proposed propellant involves 12 or more tests using charges of increasing size from 15 to 800 pounds before proceeding to a test with a 2000-pound charge. The costs for the initial tests, up to 800 pounds, range from $200,000 to $1,800,000, depending on the type of propellant. (Personal communication from Robert Geisler, A.F. Rocket Propulsion Laboratory, Edwards Air Force Base, California.) The total cost of the JANAF Tables from 1960 to 1976 was $2.3 million, less than the cost of two series of tests at the upper end of this range and 14 at the lower end. It is highly likely that without the JANAF Tables or some equivalent source of reliable thermochemical data the HBO_2 example would have been repeated many times between 1960 and 1970.

6.4 STRATOSPHERIC OZONE PROBLEM

In recent years there has been widespread concern over the extent to which the ozone concentration in the stratosphere might be reduced by the regular operation of a large number of supersonic transports and by the continued release to the atmosphere of substantial quantities of chlorofluoromethanes used in many aerosol spray cans. The Climatic Impact Committee of the National Research Council carried out extended studies on both of these problems. These studies involved mathematical modeling of 91 out of more than 200 known chemical reactions that occur in the stratosphere, and the selection of the key reactions required reliable values for the reaction rate constants for each.

These constants were obtained from a table of some 250 constants provided by the Chemical Kinetics Information Center at NBS, which was started in 1962. With their extensive files and the experience developed over this period, the Center was able to provide these critically evaluated constants in a short time. Without them, according to a member of the Panel on Atmospheric Chemistry, Frederick Kaufman (University of Pittsburgh), ". . .we would have had to gather and evaluate a huge volume of scientific data ourselves, and this large added task would have made it nearly impossible to complete the required Panel and Committee reports within the allotted time. . .Many months' work by several senior investigators and many tens of thousands of dollars would likely have been involved."

Perhaps of even greater importance is the fact that the values selected for these key constants had been thoroughly documented in the initial evaluations and were recognized as the best available. When there is widespread controversy over a proposal, such as that to ban SST's or aerosol sprays, both the advocates and opponents have a natural tendency to select the data that support their position. We saw in Chapter 3 the enormous variations that are often found in values of reaction rate constants. Without an authoritative and generally accepted set of values, it could easily become politically impossible to reach any decision requiring new legislation or regulations.

Cases of this type will probably occur with increasing frequency because of our growing concern for environmental protection, regulation of hazards, and problems of safety. Thus, data from the Chemical Kinetics Information Center, and from many other centers, will no doubt be in heavy demand in the future.

In summary, funds spent for data evaluation frequently result in savings that are greater than the cost by factors of up to several thousand. In addition, the existence of recognized sources of reliable data can save significant amounts of time and help resolve complex arguments in many cases where new legislative or regulatory proposals are under consideration. *The unsolved problem is how to allocate a portion of these widely dispersed savings to cover the costs of additional evaluation needed for the future.*

7

CURRENT ACTIVITIES AND NEEDS

7.1 ACTIVITIES OUTSIDE ORGANIZED DATA CENTERS

Although most critical data evaluation is carried out within, or with the collaboration of, an organized data center, some is done on an intermittent or occasional basis in various university and industrial laboratories. To explore one aspect of such activity, we sent a questionnaire to 30 industrial research organizations that were thought likely to carry out data evaluation in connection with their federal contract research. Ten replies, summarized in Appendix C, were received. We see that only about 20 percent of these results were published in the open literature or communicated to an established data center (with, probably, considerable overlap), and only 40 percent were reported to the sponsoring agency. Thus it appears that the results of over half of these efforts were not made available to the general scientific community in any form.

A number of data compilation activities, involving selection of data if not a full critical evaluation, are carried out or sponsored by individual companies and voluntary associations like the Copper Development Association. Industrial companies also finance the participation of their personnel in various organizations such as the American Society for Testing and Materials, which often consider material properties data in the process of drafting standards and specifications.

The above are two examples of data collected by industrial companies and the voluntary organizations that they support that might have much broader uses than those leading to their compilation. Such data might willingly be released by the companies concerned. An assessment of the extent and value of this potential source should be useful, but this would clearly require a separate study.

7.2 CURRENT STATUS OF ORGANIZED DATA CENTERS

The 37 continuing data centers listed in Appendix A operate at an annual level (FY 1977) of almost $6.8 million. The total funds from various sources are given in Table 7.1. Over 90 percent comes from agencies of the federal government. The dominant positions of the Departments of Energy and of Commerce in this listing reflect their responsibilities for nuclear data and for the operation of the National Standard Reference Data System, respectively.

TABLE 7.1 Sources of Funds for Data Evaluation Centers, FY 1977[a]

Source	Funds for Data Eval.	% of Total	Obligations for Conduct of Research & Development FY 1977				
			Basic Research	Applied Research	Total Research	Development	Total R&D
Department of Energy[b]	$3,116 K	46.8	$ 389 M	$ 498 M	$ 887 M	$ 2,688 M	$ 3,575 M
Department of Commerce	2,168	31.9	23	153	176	69	245
Department of Defense	350	5.1	294	1,476	1,770	9,118	10,889
National Science Foundation	207	3.0	625	63	688	9	697
Department of the Interior	130	1.9	129	108	237	61	298
National Aeronautics and Space Administration	106	1.6	414	791	1,205	2,472	3,677

Department of Transportation	86	1.3	---	51	51	270	321
National Institutes of Health	60	0.9	669	1,316	1,985	256	2,240
Total, federal agencies	(6,223)	(92.0)					
Other[c]	575	8.5					
Total	6,798	100.					
Total federal obligations			2,900	5,637	8,537	15,288	23,825

[a]Funds for Data Evaluation from figures from individual centers, Appendix A. Obligations for Conduct of Research and Development from Special Analysis P, Research and Development, in *Special Analyses, Budget of the United States Government, Fiscal Year 1979*, Superintendent of Documents, U.S. Government Printing Office, Washington, D.C. 20402, Stock No. 041-001-00157, plus supplemental information from National Science Foundation.
[b]Source shown in Appendix A as the Energy Research and Development Agency (ERDA), since absorbed into the Department of Energy with about 90 percent of the total DOE budget for R&D.
[c]Industrial organizations, sale of data sheets, university contributions.

It should be recognized that the research obligations listed here cover programs in many fields. For example, 90 percent of the research financed by the National Institutes of Health is in the life sciences. Almost 50 percent of the basic research of the Department of Commerce and about 80 percent of that of the Department of the Interior are in the environmental sciences. Both of these fields have data needs and activities not covered in this report. It is fair to note, however, that the National Science Foundation and the National Aeronautics and Space Administration, with about 25 and 60 percent, respectively, of their basic research budgets in the physical sciences, seem to give little attention to the compilation and evaluation of data from the research they support.

In some cases, the funds for data evaluation come from the R&D budgets shown. In others, they come from budgets for scientific and technical information or other categories not shown. The purpose of including the R&D obligations here is to demonstrate that for most of the large mission-oriented agencies the amounts devoted to data evaluation are such a small fraction of any part of the R&D budgets that their importance to the success of the whole R&D mission may easily be overlooked.

Most of the existing centers are unable to keep up with the evaluation of current publications in their fields. Thus, 32 centers with annual budgets totaling $6.2 million report that increases to a total of $11.6 million would be required to keep pace with current material. In addition, 28 centers with annual budgets totaling $4.9 million report backlogs whose elimination would require an additional $9.8 million, corresponding to an average of two years' operations. In some areas, the backlog is over five years. We estimate that about $13 million (total) is needed to overcome this backlog in existing centers. Our conclusions as to the expansion of activity needed to increase the coverage of existing centers and to cover fields with no activities at present are given in Chapter 8.

7.3 EVALUATED DATA NEEDS FOR NATIONAL PROGRAMS

Our primary concern in this section is to estimate the likely needs for evaluated data in programs of major national concern either now in progress or projected for the near future. To accomplish this we first examine the applicability of the output of the existing data centers to scientific and technical programs of various federal agencies. Table 7.2 gives a (doubtless incomplete) listing of technical programs in which the products of various data centers are used and of the agencies concerned with such programs. The multiple uses of the outputs of most centers is a telling argument for a systematic program of data evaluation.

It is much more difficult to predict the needs for data in future R&D programs. Indeed, experience has shown that particular sets of data can suddenly become important for reasons that could not have been anticipated in any normal planning process. Unexpected data needs may also arise as new fields of science emerge. However, by studying the long-range planning documents of federal agencies it is possible to identify certain broad classes of data that have a high probability of being needed, even though specific substances and properties cannot be pinpointed with certainty.

TABLE 7.2 R&D Application--Data Center Interaction Matrix

Data Center	User Application	Agency
Atomic Transition Probabilities Atomic Line Shapes and Shifts Atomic Collision Cross Section	Communication Laser R&D Nuclear fusion research Plasma physics Astronomy Space research Atmospheric and other chemistry research Theoretical chemical and physics research	DOD,NASA, DOE,NSF, EPA
Chemical Thermodynamics Center Electrolyte Data Center Texas A&M Thermodynamics Research Center Cryogenic Data Center Thermophysical Properties Research Center JANAF Thermochemical Tables Thermochemistry for Steel Making Molten Salts Data Center	Energy conservation All energy conversion systems (geothermal, solar, batteries, fuel cells, etc.) Coal liquifaction and gasification All chemical plant design and operation Heating and air conditioning Nuclear and fossil-fuel power plants (design and operation) Rocket and missile development Water-quality control Aircraft and space-vehicle design All engine design (steam, diesel, gasoline) Industrial pollution Iron and steel production Nonferrous alloy production All ore separation and reduction processes Saline water conversion Petrochemicals Petroleum refining Thermal and sound insulation Exotic power systems	DOD,DOT, DOE, EPA, NASA, HEW

TABLE 7.2 (cont.)

Data Center	User Application	Agency
High Pressure Data Center Alloy Data Center Diffusion in Metals Data Center Superconductive Materials Data Center	Magnet development Corrosion Nonferrous alloys Alloy research Materials research Electrical transmission Electronics Solid-state research Breeder reactors Accelerators	DOE, DOD, NSF
National Nuclear Data Center Gamma Ray Spectrum Catalogue Physical Data Group Nuclear Data Project Controlled Fusion Atomic Data Center Photo Nuclear Data Center Table of Isotopes Project Berkeley Particle Data Center	All nuclear reactor design Controlled fusion research Nuclear weapons systems Biological and medical research X-ray technology Radiation shielding Chemical research Theoretical and experimental nuclear-physics research Accelerator research and design Metalurgical research Nuclear regulation Geological dating Environmental protection Space exploration and research Exotic power systems	DOE,HEW, NASA,DOD, EPA
Phase Diagrams for Ceramists Electrolyte Data Center National Center for Thermodynamic Data on Minerals Rare Earth Information Center Crystal Data Center	Geological research Oil and gas mining and development Ore exploration and extraction Coatings Solid-state research Forensic chemistry Inorganic materials production Electronics Mining R&D Ceramics and glass R&D	DOI, DOE, U.S. Geological Survey

TABLE 7.2 (cont.)

Data Center	User Application	Agency
Chemical Kinetics Data Center	Coal liquifaction Coal gasification All chemical research Environmetnal pollution Exotic power systems Explosives Electronics Internal combustion engine research Battery research, electro-chemistry Rocket and torpedo propel-lant research Medical and biological research Production of heavy chemicals Pharmaceutical research Industrial pollution Agricultural pollution	EPA,DOT, DOD,NIH, NASA,DOE

TABLE 7.3 Some Data Needs in Current Federal R&D Programs and Their Current Degree of Critical Evaluation

Types of Data	Relevance to Federal R&D Programs[a]	Present Data Center Activity in the United States[b]	Level of Evaluative Activity in the United States[c]
1. Thermodynamic Data: Energy and Equilibrium			
Phase equilibria and solubility	2a,2d,2i,2j,3k,4b, 4k,4n,5a,6f,6n,7a, 7c,7d	A8,B1,B2,B3,B4,B5, B8,B10,B11,B12,B13, B14	1
Equilibrium at high pressures	1e,2d,2i,2j,4b,4c, 4f,4i,4k,5d	B7	2
Equilibrium at high tempera-tures	1e,2g,2i,2j,4b,4c, 4f,4i,4k	B9	2
Compressibilities (all states of matter)	2i,2j,4b,4k	B4,B7	2
Heat capacities and entropies	2d,2i,2j,4b,4k	B2,B4,B5,B6,B14	2
Heats of formation and combustion	2a,2c,2d,2e,2i,2j, 4b,4d,4f,4i,4k,6c	B2,B4,B6,B11,B14	2
Chemical equilibria	2d,2i,2j,3d,4d,4f, 4s,6c,6n,7a,7g	B2,B3,B4,B9,B12	2
Electrochemical (cell potentials)	2e,2i,2j,4n,4s,6c	B2,B3	1
Surface tensions and energies	2i,2j,4b,4m,6b,6c	A8	1
2. Atomic, Molecular and Crystal Structure			
Mass spectra	1c,2j,3i,4a,4b,4c, 4f,6b,6l,6m,7a,7b, 7c,7d	A3,A4	2
X-ray and uv spectra	1c,1d,1e,2j,4a,4b, 4c,4f,4g,6b,6l,6m, 7a,7b,7c,7d	A1,A2,A3,A9,A16	2
Visible spectra	1c,1d,1e,2j,4a,4b, 4c,4f,6b,6l,6m,7a, 7b,7c,7d	A1,A2,A3,A16,C1, C4	2

TABLE 7.3 (cont.)

Types of Data	Relevance to Federal R&D Programs[a]	Present Data Center Activity in the United States[b]	Level of Evaluative Activity in the United States[c]
Infrared and microwave spectra	1c,1d,1g,2j,4a,4b, 5a,6b,6l,6m,7a,7b, 7c,7d	D1	1
NMR and ESR spectra	2j,3i,4a,4b,6b,6l, 6m,7a,7b,7c,7d	A7	1
Photoelectron spectra, ESCA	1c,1d,2j,4a,4b,6b, 6c,6l,6m,7a,7b,7c	A3,A4	1
Dipole moments	4b,6b,6c		0
Atomic and molecular energy levels	2j,6b,6c	A1,D1	1
Molecular vibration frequencies	1c,1d,2j,6b,6c		0
Crystal structures	2j,4b,4n,4s,5a,6b	C1	2
Molecular structures	2j,3g,4b,4s,5a,6b, 6c	C1,D1	1
3. Dynamical Data: Rates of Physical and Chemical Changes			
Rates of gaseous chemical reactions	1g,2e,2j,4d,4f,4i, 4k,4l,4s,6c,6l,7c	A5	2
Rates of photochemical reactions	1c,1d,1e,2b,2e,3f, 4b,4d,4s,5a,6l,7a, 7b,7c,7d	A5	2
Rates of ion-molecule reactions	1c,1d,1e,4d,6k,6l, 6m	A5	2
Rates of reaction in liquid solution	2j,4s,6c,6i,7a,7b, 7c,7d	A6	1
Rates of heterogeneous reactions, catalysis	2d,2j,4s,6c,6i,6n, 7a,7b,7c,7d		0

TABLE 7.3 (cont.)

Types of Data	Relevance to Federal R&D Programs[a]	Present Data Center Activity in the United States[b]	Level of Evaluative Activity in the United States[c]
Electrode kinetics	2e,2i,4s,6c		0
Enzyme kinetics	2c,3g,3k		0
Membrane kinetics	3g,3h,3k,3l		0
Combustion kinetics	2a,2e,4f,4i,4k,4p, 6h,7c,7h	A5	1
Rates of phase changes (including evaporation)	2d,2i,2j,4b,4n,4q, 7d,7f,5d	C2	1
Biodegradation rates	2c,7a,7b,7c,7d,7f		0
Viscosity of gases and liquids	1a,2e,4b,4c,4g,4i, 4j,4k,4l,4q,6c,6h, 6l,6g	B5,B6	2
Heat conductivity	1c,2a,2d,2e,2i,4e, 4i,4k,4l,6g	B5,B6	2
Diffusivities	2d,2i,3g,4b,7c	B5,B6,C2	2
Ion mobility	3g,4b	A3,B3	2
4. Properties of Solids			
Mechanical	2d,2i,4b,4i,4k,4r, 5a,5d,6f,6g	B7	1
Electrical	2i,4b,4r,5a	B6,B15,C3	2
Magnetic	1e,2i,4b,4n,4r,5a	B15,C4	1
Optical	2i,4b,4n,4r,5a	B15	1
5. Nuclear Properties			
Neutron cross sections	2f,2g,2i,4f,4g,6a, 6f	A14,A12	3
Nuclear structure	2g,2h,2i,6a	A14,A11,A13	3
Photon interactions	2f,2g,2i,4f,4g	A9,A10,A14	2

TABLE 7.3 (cont.)

Types of Data	Relevance to Federal R&D Programs[a]	Present Data Center Activity in the United States[b]	Level of Evaluative Activity in the United States[c]
Gamma-ray spectra	2f,2g,2i,6a	A9,A11,A14,A15	3
Fundamental particles	2f,2i	D2	3

[a]From Appendix D.
[b]From Appendix A.
[c]Code for last column: 3, adequate activity; 0, no activity or at least no organized activity; 1 and 2, intermediate.

Our assessment of data needs of federal agencies for the near future was based on a study of the five-year planning documents of a few agencies with major mission responsibilities that require the use of reliable scientific data in various areas. Appendix D lists some of these areas by agency and illustrates the extremely broad range of scientific fields that are involved in the developments and programs proposed.

In many cases several of the programs listed in Appendix D will require data from the same field of science. Table 7.3 gives a sampling of various scientific fields in which data evaluation should be proceeding. The second column shows the various programs from Appendix D that will require reliable data in these fields. Column 3 shows the existing data centers with some activities in the fields listed in column 1. Column 4 gives a subjective assessment of the adequacy of such coverage. "Adequate" here means a judgment that the present coverage will probably be sufficient for anticipated needs over the next few years. It does *not* imply complete coverage of all data. Even with this definition it is only in the nuclear properties area that any topics are considered to be receiving adequate coverage at present.

The quantity of data available for evaluation varies considerably between the various scientific fields listed, as does the present level of funding. Thus no simple averaging of the ratings assigned in Table 7.3 is justified. Nevertheless, this assessment indicates that the present coverage is somewhere between one third and one half of that which we anticipate will be needed for the effective operation of various federal programs during the next five years.

8

Conclusions and Recommendations

The cost of operating the existing 37 data centers comes to just under $7 million a year. We concluded in the previous chapter that a program between two and three times the size of the present one is needed. This would represent a commitment of about $18 million a year to cover the data compilation and evaluation needs associated with present and projected federal programs. An additional $13 million (total, not yearly), representing about two years of operation at the current level, is required to catch up with the backlog in existing centers.

This conclusion is in reasonably good agreement with that reached by a National Research Council Evaluation Panel for the Office of Standard Reference Data in fiscal year 1975, which stated: "The Panel considers the amount of $15 million [per year] as the appropriate funding level for the National Standard Reference Data System program and an essential minimum with which this nation can exploit effectively its several-thousand-fold-larger annual investment in R&D."

The expansion of activities suggested here is based on the needs that we foresee to accomplish projected federal programs, and most of its cost should therefore be provided by the federal government. As demonstrated in Tables 7.2 and 7.3 the output from an individual data center is used by many different groups. In general, no single group can justify the cost of operating a center for its own exclusive use, although it is easily justified when the benefits to all users are considered.

Thus, to meet the goals set by the Federal Council on Science and Technology in 1963 and by Congress in 1968 (see Chapter 4) we RECOMMEND that the federal government increase its annual support for organized data evaluation activities to $18 million over a period of five years. We suggest as a reasonable schedule, an increase of $3 million a year for the first three years and $1 million a year in the fourth and fifth year.

Figure 8.1 shows one way in which these increases might be allocated between new activities and elimination of the existing backlog.

In general, the most effective data center operations are those maintaining close connections with active experimental programs in their fields. Most of the centers are small one or two person operations. The creation of the Office of Standard Reference Data (OSRD) at the National Bureau of Standards has helped greatly in coordinating the work and increasing the effectiveness of these small operations, in providing recognized

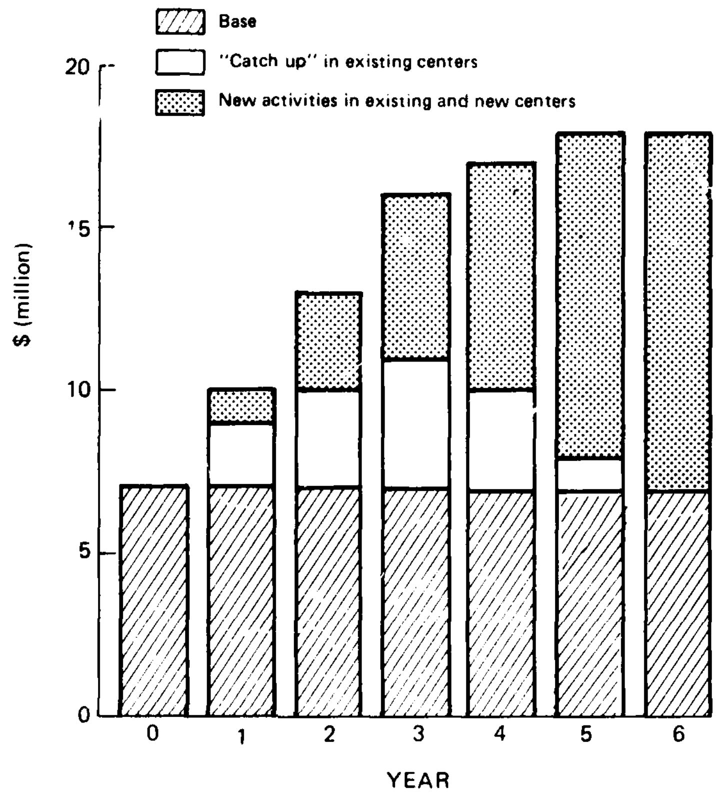

FIGURE 8.1 Possible distribution of recommended increases. This figure is intended only to suggest how the increases recommended might be used to achieve a reasonable growth of new activities while eliminating the existing backlog. The actual division of funds in any existing center would have to be worked out by that center and its sponsor(s) to meet their needs.

outlets for their results, and in guiding users to the data sources they need.

We conclude that OSRD should retain primary responsibility for overseeing all data activities and for maintaining the basic publication and dissemination program for evaluated data. However, it would be unreasonable and unsound to suggest that the Department of Commerce should therefore assume unilaterally the increased costs of an expanded program. In addition, it is clear from the experience of the National Nuclear Data Center and the JANAF Tables project that close ties between a data center and its primary customer are important.

Thus we RECOMMEND that when a particular mission relies heavily on results from a particular field of research, responsibility for data compilation and critical evaluation in that field should be assumed by the agency responsible for that mission. This should include some support for the data needs of basic science from the National Science Foundation. When more than one mission or agency is involved in a particular field, agreement should be reached on an appropriate division for the support of data activities. Areas of general interest, not primarily associated with any one mission, should remain the responsibility of the OSRD. The figures in Table 7.1 suggest that this principle is not being applied at present.

Table 7.1 also re-emphasizes the conclusion reached in Chapter 5 that the cost of critical data evaluation is a small fraction of the total federal research budget and a much smaller fraction of its total R&D budget. Under these conditions, it is difficult to ensure that data activities are not overlooked among the competing claims of much larger activities.

For this reason, we RECOMMEND that each agency be required to place its responsibility for data compilation and evaluation on one key official at a level high enough to ensure that the agency's responsibilities in this area will be fulfilled.

We saw in Section 7.1 that it is difficult to know whether data generated in the private sector and available to those who know of its existence are being utilized as widely as they should be. In addition, there is a need for a periodic review of priorities among data needs, by some group representing data generators, evaluators, and users, including representatives from industry. The Numerical Data Advisory Board (NDAB) of the National Research Council seems to be the logical group to investigate such questions, particularly if it can broaden its base of support and increase its contacts with federal agencies with emerging data needs.

Thus, we RECOMMEND that the NDAB seek to broaden its contacts with various federal agencies and accept the responsibility for periodic reviews of priorities in data evaluation.

Further, we RECOMMEND that the NDAB consider a study, or perhaps a series of studies, of the whole area of effective data management and dissemination, including the extent and general availability of data produced by private efforts.

Appendix A

TABLE A.1 Data Analysis Centers in the United States

Center	Sponsor	Funding, FY 1977
A. *Energy and Environmental Data*		
1. Atomic Energy Levels Data Center	OSRD	$ 72,000
	ERDA[a]	50,000
2. Atomic Line Shapes and Shifts	OSRD	16,000
	ERDA	40,000
3. Atomic Collision Cross Section	OSRD	145,000
	NSF	55,000
4. Ion Energetics Data Center (formerly Atomic and Molecular Ionization Processes)	OSRD	65,000
	NIH	60,000
5. Chemical Kinetics Information Center	OSRD	100,000
	DOT and NASA	72,000
	ERDA	100,000
6. Controlled Fusion Atomic Data Center	ERDA	85,000
7. Radiation Chemistry Data Center	OSRD	51,000
	ERDA	51,000
8. Molten Salts Data Center	OSRD	37,000
	NSF	30,000
9. X-Ray and Ionization Radiation Data Center	OSRD	54,000
10. Photo Nuclear Data Center	OSRD	61,000
11. Table of Isotopes Project	ERDA	200,000
12. Physical Data Group	ERDA	350,000
	NSF	35,000

TABLE A.1 (cont.)

Center	Sponsor	Funding, FY 1977
13. Nuclear Data Project	ERDA	$ 700,000
14. National Nuclear Data Center	ERDA	1,200,000
	Electric Power Research Institute	100,000
15. Gamma Ray Spectrum Catalogue	ERDA	25,000
16. Atomic Transition Probabilities Center	OSRD	45,000
	ERDA	40,000
B. Industrial Process Data		
1. Phase Diagrams for Ceramics	American Ceramic Society	3,000
2. Chemical Thermodynamics Data Center	OSRD	370,000
3. Electrolyte Data Center	OSRD	79,000
4. Texas A&M Thermodynamics Research Center	OSRD	105,000
	API	25,000
	Sale of data	160,000
	Texas A&M U.	157,000
5. Cryogenic Data Center	OSRD	105,000
	NASA	70,000
	American Gas Assoc.	32,000
6. Thermophysical Properties Research Center	OSRD	114,000
	DOD	250,000
	ERDA	100,000
	DOT	50,000
	NSF	52,000
	Payment for service	75,000
	Purdue University	70,000
7. High Pressure Data Center	OSRD	35,000
	Sale of Data	10,000
8. Alloy Data Center	OSRD	62,000
9. JANAF Thermochemical Tables	AFOSR	100,000
	ERDA	80,000
10. Data on Theoretical Metallurgy	BuMines	40,000
11. Thermochemistry for Steelmaking	Int. Copper Research Assoc.	20,000

TABLE A.1 (cont.)

Center	Sponsor	Funding, FY 1977
12. Thermodynamic Research Laboratory Data Center	OSRD	$ 90,000
13. Thermodynamic Properties of Ethylene	OSRD	70,000
14. National Center for Thermodynamic Data of Minerals	USGS	90,000
15. Electronic Properties Information Center	DSA (DOD)	150,000
	OSRD	30,000
C. Materials Utilization Data		
1. Crystal Data Center	OSRD	80,000
2. Diffusion in Metals Data Center	OSRD	24,000
3. Superconductive Materials Data Center	OSRD	9,000
4. Rare Earth Information Center	Industry	20,000
D. Physical Science Data		
1. Microwave Spectral Data Center	OSRD	56,000
2. Berkeley Particle Data Center	OSRD	22,000
	ERDA	200,000
	NSF	35,000

[a] ERDA is now part of the Department of Energy.

APPENDIX B

NATIONAL RESEARCH COUNCIL
NUMERICAL DATA ADVISORY BOARD (NDAB)
COMMITTEE ON DATA NEEDS (CODAN)
INDUSTRIAL RESEARCH INSTITUTE (IRI)
TASK GROUP ON SCIENTIFIC & TECHNICAL INFORMATION
STUDY: IMPORTANCE OF EVALUATED DATA TO RESEARCH & DEVELOPMENT

The Industrial Research Institute, Inc. (IRI), was founded in 1938, under the auspices of the National Research Council, to provide a means for coordinated study of organization and management of research and development.

Currently, the 243 member companies in the aggregate represent a major portion of the total industrial R&D effort in the United States. Member companies are from most industries committed to R&D. The numbers of professional scientists and engineers in member companies range from under thirty to over several thousand.

As part of its program to define needs particularly for validated data, the IRI Task Group for Scientific and Technical Information was asked to survey the IRI member companies via a questionnaire. Of the 243 companies belonging to IRI, to whom the questionnaires were sent, 75 responded by the deadline to meet CODAN's report publication schedule (31 percent response).

The 75 responding companies represented a number of industries. The material received was divided by industry into six broad categories, based on the major products manufactured:

1.	Industrial chemicals, energy (petroleum, coal, etc.), and metals	39%
2.	Electrical/electronic, instrument, appliances	15%
3.	Packaging/containers, paper products, coatings, films/fabrics, glass	17%
4.	Heavy equipment, steel products, rubber, transport	13%
5.	Food, drugs, pharmaceuticals, diagnostics	8%
6.	Personal, home care, detergents, toiletries	8%
		100%

Needs for numerical data in the categories listed below for use in R&D were requested. The inference can be drawn from the accompanying percentages that the use and need for data in industry is significant in all categories. The apparent exception for nuclear radiation damage, however, was found to be important in selected industries and not in others.

Categories of Numberical Data (See Exhibit I)	*% Indicating Need*
Atomic Structure (Crystallography and Defects)	61
Microstructure (Electron Microscope Level)	71
Microstructure (Optical Microscope Level)	65
Thermodynamic (Phase Equilibria, Change of State, etc.)	71
Thermal (Thermal Cond., Phonons, Diffusion, etc.)	61
Mechanical and Acoustic (Strength, Creep, Fatigue, Damping, etc.)	57
Optical (Emission, Absorption, Luminescence, Excitation, etc.)	73
Electrical (Cond., Electron Trans., Ionic Cond., Thermolec., Injection, Carrier Phen.)	44
Magnetic (Ferromagnetic, Resonance, Paramagnetic)	52
Dielectric (Ferroelectric, Breakdown, Loss, Piezoelectric, etc.)	44
Nuclear (Radiation Damage)	16
Chemical & Electrochemical (Corrosion, Battery Phen., Oxidation, Flammability, etc.)	71
Biological (Toxicity, Biodegradibility, etc.)	68
Surfact (Absorption, Surface States, Catalysis)	68

For each type of data, the following questions were asked. The average range of needs, covering all kinds of data, within an industry product category is shown in ranges in parentheses.

Do you? (See Exhibit II)

1. Use numerical data for R&D? (25-65%)
2. Make literature searches for numerical data? (25-65%)
3. Compile data books for use of your employees? (15-25%)
4. Buy commercial or government data services? (15-35%)
5. Find validated data reduces project costs? (35-50%)
6. Find validated data improves project quality significantly? (35-45%)
7. Do laboratory determination of nonproprietary materials for validation? (35-60%, in selected data areas)
8. Do laboratory determination of data of proprietary materials to produce reliable data? (30-50%, in selected data areas)

In response to the form in which the data would be useful:

71% would search for data in literature references
65% could use data bank compilations
48% could use computerized data banks

The data received apparently substantiate the use of, need for, and value of the various categories of validated numerical data. However, it is worthy to note that several R&D directors, particularly of large organizations, felt that adequate data resources do exist, that data acquisition is not a major problem, and that extensive programs in data development and validation would be difficult to justify.

Edward P. Bartkus
E. I. du Pont de Nemours & Company
Information Systems Department
July 6, 1977

Date July 6, 1977

SURVEY OF INDUSTRIAL RESEARCH INSTITUTE -
A REQUEST BY NATIONAL RESEARCH COUNCIL'S
COMMITTEE ON DATA NEEDS

Exhibit I

Number Of Companies Out Of Total Responding Indicating Needs For Each Class Of Numerical Data For R&D

	Out of 75 Companies Responding							
NEEDS FOR NUMERICAL DATA (Note that categories overlap): / Do you?	Use for R&D	Make Literature Searches	Compile Data Books	Buy Commercial/ Govt. Data Services	Find Validated Data Reduces Project Costs	Find Validated Data Improves Project Quality Significantly	Do Lab Determination of Non-Proprietary Materials for Validation	Do Lab Determination of Proprietary Materials
Atomic Structure (Crystallography and Defects)	44	36	11	27	28	30	25	36
Microstructure (Electron Microscope Level)	53	37	17	19	28	26	23	35
Microstructure (Optical Microscope Level)	49	35	16	18	24	25	26	35
Thermodynamic (Phase Equilibria, Change of State, etc.)	53	48	20	26	34	31	29	34
Thermal (Thermal Cond., Phonons, Diffusion, etc.)	46	40	11	14	24	23	22	28
Mechanical and Acoustic (Strength, Creep, Fatigue, Damping, etc.)	43	42	19	21	30	29	28	31
Optical (Emission, Absorption, Luminescence, Excitation, etc.)	55	44	20	34	37	35	36	44
Electrical (Cond., Electron Trans., Ionic Cond., Thermolec., Injection, Carrier Phen.)	33	33	10	10	21	21	15	20
Magnetic (Ferromagnetic, Resonance, Paramagnetic)	39	32	13	18	26	25	24	26
Dielectric (Ferroelectric, Breakdown, Loss, Piezoelectric, etc.)	34	25	7	9	18	14	15	19
Nuclear (Radiation Damage)	12	13	5	6	8	7	7	5
Chemical & Electrochemical (Corrosion, Battery Phen., Oxidation, Flammability, etc.)	53	53	22	27	36	36	34	38
Biological (Toxicity, Biodegradibility, etc.)	51	53	24	32	39	35	21	33
Surface (Absorption, Surface States, Catalysis)	51	48	16	23	34	30	27	35

Date June ?, 1977

SURVEY OF INDUSTRIAL RESEARCH INSTITUTE -
A REQUEST BY NATIONAL RESEARCH COUNCIL'S
COMMITTEE ON DATA NEEDS

Exhibit II

Number Of Companies In Each General Industry Category Indicating Needs For Numerical Data For R&D

Total 75 Companies

NEEDS FOR NUMERICAL DATA (Note that categories overlap): Do you use Numerical Data for R & D	29 Industrial Chemical Cos.	11 Electrical/Electronic Cos.	13 Packaging/Paper Cos.	10 Equipment/Steel Cos.	6 Food/Drug Cos.	6 Personal Care/Home Products Cos.	Total	% of Total
Atomic Structure (Crystallography and Defects)	22	7	5	6	2	2	46	61
Microstructure (Electron Microscope Level)	23	8	10	7	2	3	53	71
Microstructure (Optical Microscope Level)	20	6	10	8	1	4	49	65
Thermodynamic (Phase Equilibria; Change of State, etc.)	24	7	10	7	1	4	53	71
Thermal (Thermal Cond., Phonons, Diffusion, etc.)	18	9	8	6	1	4	46	61
Mechanical and Acoustic (Strength, Creep, Fatigue, Damping, etc.)	15	7	8	10	1	2	43	57
Optical (Emission, Absorption, Luminescence, Excitation, etc.)	24	8	8	6	4	5	55	73
Electrical (Cond., Electron Trans., Ionic Cond., Thermolec., Injection, Carrier Phen.)	13	7	6	5		2	33	44
Magnetic (Ferromagnetic, Resonance, Paramagnetic)	15	7	6	6	2	3	39	52
Dielectric (Ferroelectric, Breakdown, Loss, Piezoelectric, etc.)	14	7	6	5		2	33	44
Nuclear (Radiation Damage)	3	6		1		2	12	16
Chemical & Electrochemical (Corrosion, Battery Phen., Oxidation, Flammability, etc.)	22	6	11	9	1	4	53	71
Biological (Toxicity, Biodegradibility, etc.)	20	2	11	6	6	6	51	68
Surface (Absorption, Surface States, Catalysis)	23	8	10	4	2	4	51	68

% Companies indicating need for access to validated data: 71% References 65% Data Books 48% Computer Banks

Appendix C

NATIONAL RESEARCH COUNCIL

ASSEMBLY OF MATHEMATICAL AND PHYSICAL SCIENCES
2101 Constitution Avenue, N.W., Washington D. C. 20418

OFFICE OF CHEMISTRY AND CHEMICAL TECHNOLOGY
NUMERICAL DATA ADVISORY BOARD
Committee on Data Needs

Dear Sir:

The National Research Council Committee on Data Needs (CODAN) (list of members attached) was recently established under the aegis of the Numerical Data Advisory Board, with support from the National Science Foundation, in response to concerns from the scientific community regarding the current level of support for the compilation and evaluation of physical and chemical property data utilized in basic and applied research.

The Committee will analyze the significance of these data compilation and evaluation programs with emphasis on their contribution to major R&D programs. This analysis will include an estimate of the benefit-cost ratio of data evaluation to R&D. In addition, figures will be obtained on current Federal and private support of data compilation and evaluation in the U.S., and compared with an estimate of the data needs of major national R&D programs.

Major critical data evaluation programs are primarily carried out in certain information analysis and data evaluation centers. However, it is recognized that a considerable effort on critical data analysis may be expended as an integral part of externally funded contract R&D programs in the private sector. This critical data analysis portion of the program might be explicitly cited as a subtask of the ov[illegible] effort, or it might be accomplished under the umbrella of the usual "library [illegible]rch" associated with most R&D programs. By the critical data analysis we mean an exhaustive literature search with in-depth review of individual research publications, possible application of theoretical considerations and deriving recommended numerical values. This might involve "reinterpretation" of old data, due either to discovery of past errors, new insights, or improved procedures.

As part of the CODAN effort, it is important to estimate as accurately as reasonable the fraction of research dollars which are used for such critical data analysis. Furthermore, we wish to establish what fraction of this critical data analysis reaches the open literature. The Committee wishes to evaluate the effectiveness of this approach.

This letter is directed to a number of companies (see listing below) which carry out considerable R&D supported by government funding agencies. Through the enclosed questionnaire, we hope to obtain information of critical data evaluation efforts expended as part of this type of R&D effort. Since the information provided will at least be partially proprietary, we pledge that such information will only be seen in detail by the members of the Committee and the NRC staff involved. The Committee will only publish figures and the names of companies supplying information without associating individual companies with specific figures.

It is our understanding that your laboratory is one where critical data analysis for outside contract research projects is often carried out. We seek your aid in helping CODAN accomplish its goals by filling out and returning the enclosed questionnaire. A reply before April 1, 1977 would be appreciated.

Sincerely,

Kurt L. Wray
Member of the NAS-NRC Committee on
Data Needs (CODAN)

enclosures

Replies received from:

Aerodyne Research, Inc., Bedford, Mass.
Aeronautical Research Associates of Princeton, Inc. N.J.
Avco Everett Research Laboratory, Inc., Everett, Mass.
Bell Aerospace Textron, Buffalo, N.Y.
Calspan Corporation, Buffalo, N.Y.
Hughes Research Laboratories, Malibu, Calif.
McDonnell Douglas Research Laboratories, St. Louis, Mo.
Northrup Research and Technology Center, Hawthorne, Calif.
Physical Sciences, Inc., Woburn, Mass.
R & D Associates, Marina del Ray, Calif.

CODAN QUESTIONNAIRE

Return to: Dr. H. van Olphen
CODAN Committee
National Academy of Sciences
2101 Constitution Avenue
Washington, D.C. 20418

by April 1, 1977

1. Name of Laboratory:
2. Address: ______________________

____

____

3. Total annual research dollars expended in externally funded contract research: ______________

4. Estimate total annual research dollars used to carry out Critic Data Analysis efforts: ______________

5. Total annual number of projects requiring Critical Data Analysis efforts: ______________

6. Fraction of Critical Data Analysis efforts which are actually so designated in contractual documents: ______________

7. Fraction of Critical Data Analysis efforts which are substantially documented in final reports submitted to the responsible funding agency: ______________

8. Fraction of documented Critical Data Analysis efforts which reach the open literature: ______________

9. Fraction of Critical Data Analysis efforts which are submitted to established Data Centers: ______________

TABLE C.1 Survey of Data Evaluation Efforts Outside Formally Organized Data Centers

Laboratory[a]	A. Contract Research, $K/year	B. Critical Data Analysis Efforts $K/year	C. No. Projects/Year Requiring CDA Efforts	D. Fraction CDA Efforts so Designated	E. Fraction CDA Efforts Documented in Final Reports	F. Fraction CDA Efforts Reaching Open Literature	G. Fraction CDA Efforts to Established Center
1	13,300	19	3	1.0	1.0	0.3	0
2	0	5	1	0	1.0	1.0	0
3	1,200	150	5	0.2	1.0	0.75	0
4	3,500	350	20	0.10	0.10	0.05	0.10
5	1,113	97	14	0.36	0.57	0.14	0
6	700	150	20	0.80	0.80	0.90	0.30
7	20,000	300	30	0.20	0.80	0.05	0.02
8	7,000	50	2	1.0	1.0	0	0
9	25,000	750	50	0.30	0.10	0.02	0
10	1,200	10	1	0	0	0	0
TOTALS	73,313	1,881	146		749[b]	319[b]	86[b]
Fractions: Total E, F, G,/Total B					0.398	0.169	0.045

[a]Arbitrary order--*not* alphabetical.
[b]Sum, B · E,F,G.

Appendix D
Federal R&D Involvement by Agency

1. *National Aeronautics and Space Administration* (NASA Five-Year Planning Through 1982) Focused Activities

1a.	AERONAUTICS	Fuel economy and operating efficiency of conventional subsonic transport aircraft. Provide engines with significantly conformed efficiency, advanced aerodynamic design, improved control systems, increased use of light-weight composite materials, VTOL and STOL programs coordinated with DOD helicopter program to reduce noise and vibration.
1b.	EXPLORATION OF THE UNIVERSE	Formation and evaluation of the solar system, origin and continuing evolution of the cosmic environment. Processes by which energy is generated in the solar interior and is transformed into solar phenomena. Short- and long-term impact of solar processes on man's environment. Benefits of spaceflight to terrestrial medicine and biology.
1c.	FORMATION AND EVOLUTION OF THE SOLAR SYSTEM	Continuing study of the inner planets and the moon. Characteristics of Venusian atmosphere, gravitational harmonics, topography. Global mapping of the moon: compositional, gravitational magnetic, and heat-flow characteristics. Development of the concept of solar sailing. Jupiter-orbit probe to study pressure, temperature, density, and first-order composition of Jupiter atmosphere. Similar probes to Saturn and Uranus.
1d.	COSMIC ENVIRONMENT	Development of 2.4-meter space telescope for quantitative investigation of quasars and individual stars in nearby galaxies. Observations in infrared, ultraviolet,

and visible regions of the spectrum will be made to determine constitution, physical characteristics, and dynamics of celestial bodies and the nature of processes that occur in the extreme conditions existing in stellar objects.

1e. SOLAR TERRESTRIAL	To study and understand the physics of the processes that generate energy in the sun, transport it to earth, and couple it with the terrestrial environment. Study and understand the trigger mechanism and other physical processes in solar flares. Study interaction between solar wind and the earth's magnetosphere; study the processes that couple the magnetosphere with the ionosphere, atmosphere, and plasmas in space. Role of those constituents in the chemistry and dynamics of the lower atmosphere, beam plasma interactions, structure of the magnetospheric electric and magnetic fields, magnetosphere-ionosphere circuit-generator characteristics.
1f. LIFE SCIENCES	Human well being and performance in spaceflight. Study of life-controlling mechanisms. Application of space technology and space environment to terrestrial medical and biological problems.
1g. GLOBAL INFORMATION SERVICES	Use of observations from space in combination with ground-based analytical techniques to provide accurate continuing information concerning agricultural protection, environmental quality, natural resources, weather forecasts, and climate prediction. Includes sources classification and effects of pollutants to our water and atmosphere. Measurements of stratospheric ozone will be carried out with NOAA; measurement of radiation in space.
1h. PERMANENT OCCUPANCY OF SPACE	Development of space construction base to provide for industrialization of space: solar collectors, communication antennas, and materials research.

2. *Department of Energy* (National Plan for Energy Research)

2a.	FOSSIL FUELS	*Coal*: Direct utilization, new combustion methods, fluidized beds, boiler efficiency, stack gas technology, liquefaction, low and high Btu gas, open-cycle gas turbines, alkali metal vapor turbine, magnetohydrodynamics, *in situ* gasification, resource assessment. *Oil and Gas*: Enhanced recovery processes, resource assessment. *Oil Shale*: *In situ* combustion, gasification, hydrogenation of shale oil, separation, distillation, resource assessment.
2b.	SOLAR ENERGY	Solar electric, solar thermal, photovoltaic, wind energy, ocean thermal, advanced research and technology for novel materials, large-area silicon sheet production, heat storage systems.
2c.	BIOMASS	The development, design, construction, and operation of systems and processes for the conversion of biological materials to energy sources. The technology includes such processes as the conversion of wood or other plants to alcohol and the fermentation or decomposition of organic by-product materials to produce methane or other fuels.
2d.	GEOTHERMAL	The development, design, construction, and operation of systems and components to extract and convert the heat energy contained in geological formations, hot rocks, dry or wet steam, hot brines with associated methane, and magma. Research on scaling, corrosion, desalinization (electrodialysis), exploration technology (including seismic technology, down hole instrumentation), environmental effects of large-scale geothermal development, control of geothermal wastes, stress corrosion cracking, crevice corrosion, development of materials with improved resistance to scaling and corrosion.
2e.	CONSERVATION	Development, design, construction, and operation of buildings to minimize energy consumption, insulation, consumer appliances, heating, cooling, and ventilating systems. Advanced devices for converting

heat to electricity, fuel cells, thermionic thermoelectric, turbine systems that employ working fluids other than steam. Design and construction of electrical transport systems using extra-high-voltage, ac-dc underground and cryogenic systems. Design and construction of electrical propulsion of vehicles, ac-dc conversion equipment. Hydrogen enrichment of natural gas, battery research, electric utility load leveling (compressed air storage, underground pumped hydroelectric and thermal energy storage.) Industrial unit operation-combustion efficiencies, stack heat loss, detailed analyses for heat balances. Transportation energy conservation-engine design and operation, vehicle design to limit drag. Energy conversion systems, steam Rankine cycles, closed Brayton gas turbine, ultra-high-temperature conversion machines (materials research).

2f. FUSION POWER

Magnetic fusion-magnetic mirror system, ORMAK, TOKAMAK, and the ALCATOR; applied plasma physics research; completion of toroidal plasma device. Laser fusion, theoretical research and modeling, development of diagnostic instrumentation multiple-beam ion pulses.

2g. FISSION POWER

Liquid metal fast breeder reactors: core damage, accident containment, attenuation of radiological products, advanced field research to develop fuels with low creep and swelling, advanced carbide and nitride fuels to develop higher breeder gains.

Water-cooled breeder: develop capability of breeding in pressurized water, study physics, thermal, and fuel performances, use of thorium cycle, establish thorium and fuel recycle capability.

Gas-cooled reactor: development of high-temperature and very-high-temperature gas-cooled reactors, development of large-scale heat exchangers, turbomachinery and valves, safety and environmental studies.

Light-water reactors: hydrologic impact of thermal and radionuclide release, effect of release of atmospheric heat, improved safety systems, fuel recycling.

2h. BIOMEDICAL AND ENVIRONMENTAL RESEARCH

Develop rapid biological and biochemical automated cytochemical screening techniques, study potential genetic and developmental effects of energy production by studies of mutageneses and teratogeneses. Assess impact of surface coal and uranium mines, offshore oil, and gas development, oil toxicity, effects of fossil-fuel combustion products.

2i. BASIC ENERGY SCIENCES

Materials sciences: electronic, magnetic, optical, and thermal properties of pure materials and alloys; surface phenomena; phase transformations; stability, materials interactions; defects; diffusion and radiation effects; thermodynamics and electrochemistry; low-temperature research; superconductivity; mechanical properties.

Nuclear sciences: research relative to fusion and fusion reactors, waste management, safeguards, weapons, biomedical and environmental problems. Properties, structure, and interaction of nuclear matter, theoretical nuclear research, rare elements, enriched isotopes, develop basic chemical, physical, and nuclear data for actinide element waste disposal.

Fundamental nuclear research (high and low energy): super HILAC/Bevalac facility, Holifield heavy-ion facility, Anderson Meson Physics Facility, Bates linac high-resolution spectrometer, accelerator design and development.

2j. MOLECULAR, MATHEMATICAL, AND GEOSCIENCES

Develop basic understanding in molecular, ionic, atomic processes pertinent to all energy development programs. Chemical structure reaction mechanisms, catalysis. Support engineering sciences programs to improve technology transfer, laboratory-scale demonstration of new energy-related technologies.

3. *National Institutes of Health* (Forward Plan 1978-1982)

3a. AGING

Studies on basic aging process with emphasis on biological phenomena and age-related disease; cellular, biochemical, nutritional, immunological, physiological processes; metabolism of therapeutic drugs; pathology.

3b.	ALLERGY AND INFECTION	Immune system, infections, vaccines, antibiotics, diagnostic methodology, chemotherapy, nucleic acid recombinants.
3c.	METABOLISM AND DIGESTIVE DISEASES	Mechanism of action of insulin, endocrinology, intermediate metabolites. Role of minerals, trace elements, vitamins, protein, amino acids, fats, fatty acids, carbohydrates in normal and disordered states. Basic mechanism of kidney function, renal dialysis technology, anemia, blood-clotting mechanism, hemophilia. Development of automated blood glucose control devices. Basic nutrition studies, radioassay of metabolites, use of high-pressure liquid chromatography, mass spectra.
3d.	DENTAL RESEARCH	Chemical action involved in dental caries, antiplaque agents, plaque enzymes. Fluoride biochemistry and physiology. Materials research, development of less corrodible amalgams, tooth implants, dental prosthetics.
3e.	ENVIRONMENTAL HEALTH	Mechanism of toxicity of environmental agents, absorption mechanisms. Identity of toxic agents, identification and determination of toxic substances in air, water, food, and occupational situations and marine environments.
3f.	DISEASES OF THE EYE	Chemistry of visual pigment, electrical activity of photoreceptor cells, [illegible]-laser development
3g.	MEDICAL SCIENCES	Biophysics of cell structure, biochemical and organic chemical studies relative to cell structure and metabolism, [illegible] and characterization of proteins, transport across membranes, ion [illegible], structural relationship of [illegible] biological activity, factors [illegible] drug metabolism.
3h.	NEUROLOGICAL AND COMMUNICATIVE DISORDERS	Development of prosthetic devices [illegible] deaf, blind, paralyzed. Effects of [illegible] and noise on hearing, neurochemistry, lipid metabolism, chemical synthesis of neuromal membranes, molecular biology of nervous system.
3i.	RESEARCH RESOURCES	Use of digital computation, mass spectroscopy, nuclear magnetic resonance, electron

microscopy, computer-based technical information center for research support, activation analysis.

3j.	NATIONAL LIBRARY OF MEDICINE	Development of information-transfer systems, use of satellites for information communication, information storage and retrieval, application of laser and fiber optics to provide interference-free large-band channels.
3k.	HEART AND LUNG	Chemical effects of interaction of blood components with foreign materials, drug metabolism, standardization of drugs, separation and purification of blood components, development of blood substitutes. Research on enzyme chemistry, rapid estimation of enzyme concentrations.
3l.	CANCER	Identity of cancer-causing chemicals, methods of analysis, metabolism of toxic agents, mathematical modeling of cancer causes and treatment effects, nutrition studies, use of ultrasonics, development of antitumor agents, metabolism of antitumor agents.

4. *Department of Defense*

AIR FORCE

4a.	HUMAN RESOURCES AND LIFE SCIENCES	Detection and identification of pollutants and other chemicals, detection of electromagnetic radiation, develop precise instrumental methods of analysis (mass spectroscopy, laser technology), prediction of environmental impact.
4b.	MATERIALS	(a) *Structural materials*: development of high-temperature materials resistant to corrosion, thermodynamically stable. Fundamental research in mechanical properties of materials, phase-transformation kinetics, metastability, relationship between microstructure and physical properties, intermetallic compounds, development of ceramics and glasses with improved strength, creep resistance, thermal fatigue and processability, titanium-aluminum alloys, joining phenomena, surface behavior, mechanism of brittle fracture, plastic deformation, corrosion and stress corrosion, composites and polymers, fundamentals of metalworking.

(b) *Environmental-resistant materials*: radiation-resistant materials, erosion and corrosion resistance, oxidation mechanisms, radiation damage mechanisms, material coatings, interfacial reactions, diffusion.

(c) *Electromagnetic materials*: development of new classes of lasers, optical coating materials, infrared detecting materials, fiber-optics materials, magnetic bubble materials, new electronic materials (organic materials with transition metal complexes), studies on semiconductor-dielectric interfaces, superconducting materials, acoustic and optical materials, theoretical studies on electromagnetic materials.

(d, *Fluids, lubricants, and containment materials*: synthesis and characterization of new elastomeric materials, additives for high-temperature lubricants.

4c. GEOPHYSICS

Propagation of electromagnetic radiation, effect of sun radiation, interplanetary plasma, magnetosphere and ionosphere and their interactions. Optical-infrared spectroscopy and sensors, atmospheric absorption and scattering, theoretical studies, nonequilibrium radioactive phenomena, lasers. Gravity measurements, noise and motions in earth's crust, seismological and geological studies.

4d. ENVIRONMENT

Physical and chemical properties of the upper atmosphere, atmospheric composition and chemistry, upper-atmosphere predictions.

4e. AEROSPACE VEHICLES

Structures and structure dynamics, aerodynamic turbulent boundary-layer flows, aerophysics, aerodynamic noise, missile dynamics, heat transfer, trunk flutter, mathematical modeling of systems design.

4f. PROPULSION

Rocket propulsion, solid propellants, amine propellants, combustion kinetics and mechanics, thermophysical properties and rocket systems, plasma propulsion, fluid mechanics, fuels, pollutants, emission and control noise, lubrication. Batteries, solar cells, MHD power generators, MHD lasers.

4g. WEAPONRY

Nuclear-weapons effects, chemical effects in lower atmosphere, chemical effects in upper atmosphere, thermal radiation effects, shock waves, electromagnetic pulses, radiation effects in detector materials and satellites. Conventional weapons, subsonic and unsteady aerodynamics, missile guidance and control, materials. Electromagnetic weapons, gas-dynamic lasers, electric discharge lasers, chemical lasers, charged-particle beams.

4h. ELECTRONICS

Image sensors, antenna design and structures, radar design and detection, microwave propagation and detection, nuclear magnetic resonance gyroscopes, geodesy, ring lasers, surface acoustic-wave modulation and detection, time standards, processing and transmittance of image information, shortwave transmission, magnetic-wave devices, high-power microwave tubes, electro-optical devices, microcircuit research, theoretical research on electromagnetic materials. Laser developments, near-infrared tunable lasers, short- and ultra-short-pulsed lasers, holography, nonlinear optics, propagation of laser pulses, physics of high-energy gas lasers. Automatic speech recognition, vehicle signatures. Computer software, hardware, and signal processing. Metal insulator semiconductors, charge-coupled devices, radiation-induced transport phenomena.

ARMY

4i. BALLISTICS

[illegible] kinetics of propellants and explosives, heat transfer and erosion, ignition and flame dynamics, turbulent boundary-layer development, projectile aerodynamics, shaped charges, muzzle gas flow. Magnus effect on supersonic projectiles, ballistic response of materials, structural response to dynamic blast-pressure loading.

4j. AIR MOBILITY

Aerodynamics mechanisms contributing to dynamic loads of rotors, rotor blade dynamics, stall, control system reliability, blade acoustics, variable-blade geometries, rotary-wing dynamic research, flap-lag-torsion stability of elastic blades. Structural and aerodynamic design. Research

		on gears, bearings, seals, lubrication, powder metallurgy, composite materials.
4k.	MUNITIONS	Combustion, ignition, detonation of explosives, propellants, pyrotechnics. Ultra-high-pressure physics. Energetic material storage. Reaction mechanisms, coatings, adhesion, ultra-high-purity alkali halides, fracture mechanics, superconductivity, structural analysis. Synthesis of explosives, detection and analysis of explosive materials in liquid and solid media, effects of heating rate on pyrolysis of explosives, role of free radicals in modifying initiation thresholds, fuel-air explosive detonation studies.
4l.	BASIC SCIENCES AND ENGINEERING	Atmospheric sensing and probing, cloud and aerosol physics, biodegradation of materials, chemical and biological weapons defense, chemistry of surfaces and interfaces, atmospheric chemistry, electronic materials, antennas and detection of radiation, signal processing, man-machine interfacing. Applied mathematical analysis for heat-transfer studies, statistical techniques for field data, operations research. Solid mechanics, fluid mechanics, engines and fuels. Ceramics, polymeric and metallic materials, fundamental physics research. Nonlinear chemical reactions.
4m.	COMMUNICATIONS AND ELECTRONICS	Physical electronics, electron devices, antennas and electromagnetic detection, circuits, networks, signal processing, information processing. Surfaces and interfaces in solid-state electronics, high-power pulsed radar, computer-aided design of hybrid integrated circuits, process modeling and simulation.
4n.	MATERIALS	Research in corrosion, oxidation, radiation, decomposition. Effects of structure, defects, and composition on physical and chemical properties. Chemical composition and microstructure of special alloy steels. New synthetic methods. New concepts in testing and analysis of materials (optical probing to detect fatigue, magnetic-field interactions).
4p.	MATHEMATICS	Applied analysis, numerical analysis, operations research, statistics and

probability, computer science as applied to aerodynamics, heat transfer, structural analysis, communication, chemical kinetics and combustion, guidance and control missiles.

4q.	MECHANICS AND AERONAUTICS	Solid mechanics, fatigue and fracture, shock, vibration, wave propagation, surface mechanics, composites, shock loads, noise source and abatement, lubrication, friction, wear. Fluid mechanics, aerodynamics, aeroacoustics, ballistics, missile aerodynamics, rotor-generated noise. Fuel conservation, propellants.
4r.	PHYSICS	Laser research, crystal growing, miniaturization of laser range finders, rare-earth lasers, frequency conversion, thermal imaging. Precision navigation, electric discharge phenomena, structure of solids, crystal defects, electronic and non-electronic transport properties, surface and interface phenomena, dielectric properties of materials, photoelectric and optoelectronic devices and systems. Hybrid and monolithic charge-coupled imagers, uncoated thermal imaging concepts, photocathode materials, III-V charge-coupled devices, py electric vidicons.
4s.	CHEMISTRY	Polymer chemistry, high-energy materials, photochemistry, chemiluminescence, sensing and detection of chemical agents. Photodegradation, atmospheric chemistry. Aerosols, microemulsions, chemical lasers, electrochemical energy conversion.

5. *Department of Defense*

ARPA

5a.	MATERIALS	Solid brushes for collection of current in high-power density, segmented magnet electrical machines, rapidly solidified powders, superalloys for turbine blades and aircraft structures, ultra-high-strength high-carbon steels, wear theory and monitoring analysis, metal matrix composites. Advanced optical ceramics for all-weather, high-Mach, multimode, electromagnetic windows. Corrosion-resistant coatings using laser processing, infrared laser windows, low-heat-loss nozzles. Photo-

cathodes, pyroelectric and ferroelectric materials for acoustic transducers and uncooled thermal imagers, silicon infrared detectors, integrated-circuit design and processing, large-scale integrated circuits, III-V compound semiconductors.

5b.	CYBERNETICS	Learning strategies, self-assessment, transformation of written material to visual images, skill acquisition using neural man-made links, automated group decisions, spatial information storage and retrieval system, ultra-rapid prose/picture presentation, aids to reasoning, problem solving, heuristic modeling, motor skills.
5c.	COMPUTER AND COMMUNICATION SCIENCE	Automated cartography, intelligent information processing, automated Morse code operator, complex laser operations, information overload systems. Development of memory bits in the 10^{15}-10^{17} range, archival memory technology, interdependent network systems, user support systems, distributed file structures, intelligent data base systems, advanced terminals, advanced network systems. Basic machine intelligence, applied machine intelligence, natural language research, advanced digital structures, real-time symbolic processing.
5d.	GEOPHYSICS	Tunnel (voids) locating using passive and active sensing techniques (electromagnetic, acoustic, gravity) using airborne, surface, and subsurface methods. Stress-wave propagation, plastic yield effects, planar and curved shock fronts, effects of pre-existing stress fields, heterogeneous materials with nonplanar boundaries, anisotropy, free surface effects, interaction of nonlinear shock waves, cracks and faults in materials.

6. *Department of Defense*

NAVY

6a.	NUCLEAR PHYSICS	Radiation damage, detection, shielding, solar and cosmic rays, helium embrittlement and creep, dosimetry, radiation damage to satellites and other communications systems, magnetic detection systems. Development of alloys for high-temperature high-flux nuclear reactors, nuclear

activation techniques for improving photographic images. Ion implantation techniques, use of electron beams for curing adhesives and composites. Charged-particle beams and beam propagation.

6b. GENERAL PHYSICS	Basic research on electronic, magnetic, optical, structural, and thermal properties of materials. Surface and interface physics, crystal defects, high-electric-field carrier transport. Atomic and molecular properties relevant to communications, lasers, chemical lasers, chemical synthesis, directed energy systems. Laser stability, high-power blue-green lasers, optical properties of the atmosphere, molecular collisions in intense optical fields, nonlinear optical effects, integrated optical microcircuits, relatavistic electron-beam propagation. Physical acoustics, behavior of sound in nonlinear media, nondestructive identification of microscopic stresses and failure mechanisms in solids, interaction of acoustic waves with submerged objects. Interaction of ions, electrons, atoms, molecules, photons in ionosphere blackout; MHD; gigawatt microwave sources; communication enhancement; gaseous discharge devices; laser systems; lightning breakdown. Superconductivity, superconducting junctions and arrays, logic switching, information storage, multifilament superconducting wires.
6c. CHEMISTRY	Kinetics of chemical lasers, interfaces and surfaces, analysis of surfaces, ESR and NMR techniques, adhesion, lubrication, determination of trace metals in seawater, engine oils, other fluids, synthesis of environmentally stable polymeric materials, caborane, siloxane, and phosphazene elastomers, coatings, fluoroepoxies, polymers of useful electrical and pryoelectrical properties. Fuel cells, batteries, catalysis, electrochemical reactions. Inorganic polymers, ceramics, composites, piezoelectric materials.
6d. MATHEMATICAL SCIENCES	Mathematical research leading to acquisition and analysis of data, analyses of fluid flow in ship and missile design. Communication systems, structural analysis,

information processing, storage, and retrieval. Computing and information processing systems and devices. Numerical analysis, statistical modeling and analysis, digital computer simulation, applied mathematics and control theory, logistics, operations research.

6e. ELECTRONICS

Electromagnetic-wave propagation and radiation, reflection, refraction, scattering; antenna theory and radar target detection. Physical electronics, electronic materials, semiconductors, other electronic materials, surfaces, and interfaces. Electronic components, microwave- and millimeter-wave devices, ion implantation, defects and radiation effects in solids, integrated circuits, signal sources, radiation detectors, circuit and control theory, network analysis, linear and nonlinear system theory, distributed processing, signal coding, signal processing, and fault analysis.

6f. MATERIALS

Metals and alloys, laser surface treatment, amorphous metals, permanent magnets, bubble memory, laser welding of titanium alloys, fiber reinforcement for high-temperature composite materials, sustained load cracking in titanium alloys, niobium-titanium superconductors, fatigue, fracture and environmental effects. Effects of nucleators for grain growth, superalloys, hybrid composites. Corrosion, environmentally assessed fracture of high-strength alloys, hydrogen embrittlement, stress corrosion resistance, hydrogen-metal reactions, corrosion fatigue-failure mechanisms, cathodic protection, corrosion behavior of amorphous alloys. Ceramics, seal materials, nozzles, sonar ceramics, vapor-phase processing of ceramics, new composite concepts for ceramics, role of surfaces in degradation of optical properties of insulators. Radiation resistant materials.

6g. MECHANICS

Hydrodynamics, numerical computation of free surface flows, boundary-layer heating, drag reduction, interaction of internal waves and free surface, gas-lubricated bearings, extreme motions of ships, counterrotating propellers. Aerodynamics, aircraft and missile configurations, effect of body temperature on boundary-layer transition at supersonic speeds, calculation of compressible boundary-layer flow, minimum wave drag configurations, high subsonic airfoils. Structural mechanics, radiation of sound from submerged structures, structural response of submarines to high-intensity time-dependent coupled thermomechanical loading. Failure criteria for composites, adhesive joints, metals and brittle materials. Optimization techniques for determining critical wing and external stores flutter configurations, influence of structural nonlinearities on aeroelastic stability of control surfaces.

6h. ENERGY CONVERSION

MHD development and research, design and construction of a 1-megawatt (thermal) experimental generator. Turbine and rocket engine combustion, safe storage of high-energy missile propellants and explosives, combustion kinetics, turbine and engine efficiency, transonic turbines, electrical high-power density engines, wear and lubrication studies for engine performance.

6i. OCEANOGRAPHY

Acoustic research, ocean environment, ocean currents, surface and mixed layer, internal waves, sea floor, sediment reflectivity and attenuation. Remote sensing using satellites and radar. Airborne magnetic anomaly detection. Water pollution, fouling, corrosion, disposal of dredge, spoils. Improved sea/swell weather forecasting.

6j. TERRESTRIAL SCIENCES

Geography, detection of surface currents, nearshore and on-bottom features induced by sediment transport, develop master prediction model of coastline, research on over-the-horizon radar, meteorological satellites, wave statistics. Arctic research, long-range arctic environmental and acoustical forecasting, mass-energy exchange between arctic basin and peripheral seas, ice-dynamics modeling, remote sensing of magnetic field of arctic basin.

6k.	EARTH PHYSICS	Nature and distribution of earth's crust, geothermal and earthquake prediction, tidal prediction, environmental effects. Ionospheric irregularities, effect of D-region ionosphere nonlinearities on radio communications, theoretical studies of spread-F and field-aligned ionospheric irregularities.
6l.	ATMOSPHERIC SCIENCES	Lower atmosphere and marine boundary layer, marine fog and aerosol distribution, effects on optical and electromagnetic transmission, cloud physics. Upper atmosphere, ionospheric plasma dynamics, solar control of the ionosphere and atmosphere, measurement techniques and instrumentation, remote sensing, geometry and spectral content of solar phenomena, solar flares.
6m.	ASTRONOMY AND ASTROPHYSICS	Celestial radio sources, precise astronomical position reference systems, radio interferometer techniques, determination of atmospheric constituents, measurement of space emissions, galactic x-ray and gamma-ray measurement, analysis of fluxes and energies of cosmic rays (heavy particles), enhanced solar disturbance predictive capability.
6n.	BIOLOGICAL AND MEDICAL SCIENCES	Physiology, diving effects, decompression, gas mixtures, high-pressure neurological effects, frozen blood components, motion sickness, high-pressure deafness, oxygen toxicity, effects of low-frequency and middle-frequency electromagnetic fields on various biomedical, biophysical, and biochemical parameters. Biochemical events of injury and healing. Marine biological processes destructive to wood and metal.

7. *Department of the Interior, Environmental Protection Agency, Department of Transportation* (Near Term)

7a.	TOXIC SUBSTANCES	Identification and characterization of toxic substances (joint with NIH). Chemical determination in waters, air, development of standards, chemical decomposition rates in waste water, metabolism, transport properties, agricultural effects, ecological effects.

7b.	PESTICIDES	Toxicity, chemical determination in groundwater, water drainage systems, decomposition rates and products in air and water, ecological effects, biochemical studies, storage, hydrologic research, marine biology.
7c.	AIR POLLUTION	Chemistry of the atmosphere, chemical reactions in troposphere and stratosphere, identification and source of pollutants. Toxicology, water solubility, diffusion rates of pollutants, engine combustion kinetics, fuels, photochemistry of all chemical components in upper atmosphere, heterogeneous reactions, development of new and more sensitive techniques for chemical reaction rate measurement, molecular spectroscopy, eddy diffusion models with time-dependent coupled transport and chemical kinetics, radiation and thermal transfer mechanisms.
7d.	WATER POLLUTION	Systematic study of water compositions, contaminents, aqueous oxidation-reduction, sources of contaminants, waste-water treatment, mine effluents, sewage disposal and treatment, vapor-pressure data, volatilization rates, reaction rates in water and saline water. Agricultural wastes, hydrology, marine life, groundwater contamination.
7e.	WATER RESOURCE	Geological survey, digital cartography, hydrologic studies of nuclear landfill disposal, hydrologic research, coal resource development, water quality, mine waste runoff, earth-science studies, remote-sensed satellite data, remote-sensed imagery data, desalinization, demineralization.
7f.	HAZARDOUS MATERIALS	Identification of hazardous materials, thermodynamic stability, volatilization data, vapor pressures, photochemical reaction rates, pipeline transport-hydrogen stress-corrosion cracking, hydrogen enbrittlement, fatigue, corrosion.
7g.	MINING AND METALLURGY	Geological data, research on process metallurgy, analysis for trace elements, metallurgical process waste disposal, chemical reaction rates, energy conservation in metallurgical processes.

Stack gas desulfurization, recovery of minerals from wastes, deep sea bed minerals, oil shale mining.

7h. TRANSPORTATION	Fuel efficiency, emission control, climatic impact of stratosphere flight, noise abatement, fire safety, development of quality-control procedures, remote measurement of ice characteristics, remote sensing and detection of oil spills, marine environment, waste-water treatment, solid-waste disposal. Satellite navigation aids and detection systems, communications, computer design. Effects of seismic and wind motions on bridges, tunnels. Soil mechanics, environmental effects on soils, remote-sensing techniques, automated signal equipment, advanced propulsion systems, regenerative braking, automated traffic control.